2964

RECEIVED
STIC
BIOTECH GRP

ONLINE SEARCH REQUEST FORM

* *

'95 SEP 22

USER _Spector_   SERIAL NUMBER _386680_

U.S. PAT. & TM. OFF.

ART UNIT _1812_   PHONE _308-1793_   DATE _9/22/95_

Please give a detailed statement of requirements. Describe as specifically as possible the subject matter to be searched. Define any terms that may have special meaning. Give examples or relevant citations, authors, or keywords, if known.

You may include a copy of the broadest and or relevant claim(s).

Please search commercial + interference
& issued databases for
        Seq ID No: 1 of 08/167628 (parent)

If possible, do search of EST's & STS's
separately (not lumped in with everything
else) - it obscures results if all done
                    at once
            Thanks

        4/24/95 CRFA seq moved
                        to public dbase

        PN = 5,408,040

* *

```
KW   stringent hybridisation; ss.
PN   WO9300430-A.
PD   07-JAN-1993.
PF   25-JUN-1992; F00589.
PR   25-JUN-1991; FR-007807.
PA   (CNRS ) CENT NAT RECH SCI.
PI   Martinerie C, Perbal B;
DR   WPI; 93-036377/04.
PT   Nucleotide sequences hybridising to regions of chicken nov gene -
PT   useful as probes for detecting complementary sequences to
PT   evaluate development and/or differentiation of tumours
PS   Claim 14; Page 34; 67pp; French.
CC   This sequence is homologous to exon 4 of the CTGF gene and was used
CC   as a probe in Northern hybridisations with different human tissues
CC   and cell lines. Probes homologous to exon 2 of the chicken nov gene
CC   and to exon 3 of the CTGF gene (Q36035 and Q36039, respectively)
CC   were also used in the hybridisation experiments. The results showed
CC   that the human homologue of the chicken nov gene and the CTGF gene
CC   belonging to the same family are expressed differently in different
CC   tissues or lines as different RNA transcripts.
SQ   Sequence   435 BP;   120 A;   106 C;   114 G;   95 T;

DB 6;   Score    65;   Match 67.2%;   QryMatch 3.1%;   Pred. No. 2.43e-19;
Matches   127;   Conservative    0;   Mismatches   62;   Indels    0;   Gaps   0;

Db  104 gcttacaggccagagccacccctaggagtagaagtctctgactcaagtgtcaactgcatt 163
        |||||| | |||| ||| ||| | ||   ||    ||  | |||||||||||||||||
Qy  670 GCTTACCGACTGGAAGACACGTTTGGCCCAGACCCAACTATGATTAGCCAACTGCCTG 729

Db  164 gaacagaccacagagtggacagcatgctccaagagctgtggtatgggttctccacccgg 223
        | ||||||||||||||||| || |||  ||||||||| |||| |||||||||||||||||
Qy  730 GTCCAGACCACCACAGAGTGGAGCGCCGTGTTCCAAGACCTGTGGGATGGGCATCTCCACCCGG 789

Db  224 gtcaccaataggaaccgtcaatgtgagatgctgaaacagactcggctctgcatggtgcgg 283
        || ||||||   |   |  | ||   |||||| |||||||||  |||| |||||||| |
Qy  790 GTTACCAATGACACACGCCTCCTGCAGGCTAGAGAAGCAGAGCCGCCTGTGCATGGTCAGG 849

Db  284 ccctgtgaa 292
        || || |||
Qy  850 CCTTGCAA 858

RESULT  15
ID   Q36041 standard; cDNA; 216 BP.
AC   Q36041;
DT   24-MAY-1993  (first entry)
DE   Homologous to chicken nov gene exon 4.
KW   avian nephroblastoma; avian myeloblastoma virus;
KW   stringent hybridisation; ss.
OS   Homo sapiens.
PN   WO9300430-A.
PD   07-JAN-1993.
PF   25-JUN-1992; F00589.
PR   25-JUN-1991; FR-007807.
PA   (CNRS ) CENT NAT RECH SCI.
PI   Martinerie C, Perbal B;
DR   WPI; 93-036377/04.
PT   Nucleotide sequences hybridising to regions of chicken nov gene -
PT   useful as probes for detecting complementary sequences to
PT   evaluate development and/or differentiation of tumours
PS   Disclosure; Page 36; 67pp; French.
CC   Nucleotide sequences which hybridise to Fragment XIII (Q36038) of
CC   the 4th. exon of chicken nov gene under stringent conditions (i.e.
CC   50% formamide, 5 x SCC) are claimed. The claimed sequences preferably
```

```
CC   encode a protein with at least 85% homology to R31604 which is encoded
CC   by Fragment XIII. Such nucleotide sequences are further
CC   characterised in that they comprise at least part of a 400bp HincII
CC   fragment derived from a recombinant clone isolated from a human
CC   placenta DNA library. Sequence XVII (Q36041) is a subfragment of
CC   sequence XVI (Q36040) but is not specifically mentioned in the
CC   claims.
SQ   Sequence   216 BP;   62 A;   56 C;   60 G;   38 T;

DB 6;   Score    65;   Match 67.2%;   QryMatch 3.1%;   Pred. No. 2.43e-19;
Matches   127;   Conservative    0;   Mismatches   62;   Indels    0;   Gaps   0;

Db    1 gcttacaggccagagccaccctaggagtagaagtctctgactcaagtgtcaactgcatt 60
        |||||| | |||| ||| |||  | ||   ||   | |||||||||||||||||
Qy  670 GCTTACCGACTGGAAGACACGTTTGGCCCAGACCCAACTATGATTAGCCAACTGCCTG 729

Db   61 gaacagaccacagagtggacagcatgctccaagagctgtggtatgggttctccacccgg 120
        | ||||||||||||||||| || |||  ||||||||| |||| |||||||||||||||||
Qy  730 GTCCAGACCACCACAGAGTGGAGCGCCGTGTTCCAAGACCTGTGGGATGGGCATCTCCACCCGG 789

Db  121 gtcaccaataggaaccgtcaatgtgagatgctgaaacagactcggctctgcatggtgcgg 180
        || ||||||   |   |  | ||   |||||| |||||||||  |||| |||||||| |
Qy  790 GTTACCAATGACACACGCCTCCTGCAGGCTAGAGAAGCAGAGCCGCCTGTGCATGGTCAGG 849

Db  181 ccctgtgaa 189
        || || |||
Qy  850 CCTTGCAA 858

Search completed: (Fri Sep 22 22:19:45 1995
Job time : 130 secs)
```
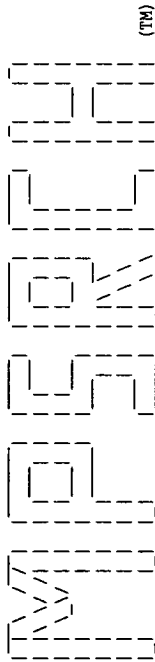
```
*****************************************************

   [WORM ASCII ART LOGO] (TM)

*****************************************************

          Release 2.1D John F. Collins, Biocomputing Research Unit.
          Copyright (c) 1993, 1994, 1995  University of Edinburgh, U.K.
                    Distribution rights by IntelliGenetics, Inc.

*****************************************************
```

MPsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Fri Sep 22 22:17:35 1995;   MasPar time 117.89 Seconds
                                    898.163 Million cell updates/sec

Tabular output not generated.

Title: >US-08-167-628-1
Description: (1:2075) from 5408040.seq
Perfect Score: 2075
N.A. Sequence: 1 CCCGGCCGCACAGCCCCGAGA.........AGGAAATGTGGTAGCTCACG 2075
      Comp: GGGCCGGCGTGTCGGGGCTCT.........TCCTTTACACCATCGAGTGC

Scoring table: TABLE default
                Gap 6

Nmatch STD :   Dbase 0; Query 0

Searched: 61539 seqs, 25515148 bases x 2

Database: n-geneseq
                 1 n-gen1
                 2 n-gen2
                 3 n-gen3
                 4 n-gen4
                 5 n-gen5
                 6 n-gen6
                 7 n-gen7
                 8 n-gen8
                 9 n-gen9
                10 n-gen10
                11 n-gen11

Statistics: Mean 10.059; Variance 7.188; scale 1.400

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                        SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	921	44.4	2330	4	Q26422	Gene for beta-IG-M2.	0.00e+00
2	264	12.7	1975	6	Q36031	Chicken nov coding se	2.19e-132
3	240	11.6	619	6	Q36046	Human sequence XXVII	2.44e-118
4	215	10.4	684	6	Q36032	Chicken nov gene frag	9.12e-104

Result No.	Score	Query Match	Length	DB	ID	c	Description	Pred. No.
5	160	7.7	461	6	Q36044		Chicken nov gene exon	5.23e-72
6	106	5.1	252	6	Q36035		Chicken nov gene exon	1.64e-41
7	106	5.1	196	6	Q36045		Chicken nov gene exon	1.64e-41
8	103	5.0	131	6	Q36049		Fragment XXXII encode	7.54e-40
9	88	4.2	2028	4	Q26421		Gene for beta-IG-M1.	1.28e-31
10	80	3.9	1047	2	Q10572	c	Human Natriuretic Pep	2.72e-27
11	79	3.8	1047	2	Q10572		Human Natriuretic Pep	9.36e-27
12	67	3.2	372	10	Q57417		CYR61 like protein.	2.18e-20
13	65	3.1	216	6	Q36039		Homologous to exon 3	2.43e-19
14	65	3.1	435	6	Q36040		Homologous to exon 4	2.43e-19
15	65	3.1	216	6	Q36041		Homologous to chicken	2.43e-19
16	64	3.1	230	6	Q36033		Chicken nov gene exon	8.05e-19
17	63	3.0	209	6	Q36038		Chicken nov gene exon	2.66e-18
18	62	3.0	452	6	Q36037		Sequence XI homologou	8.78e-18
19	62	3.0	253	6	Q36036		Encodes polypeptide h	8.78e-18
20	60	2.9	72	6	Q36047		Human sequence XXVIII	9.45e-17
21	54	2.6	142	6	Q36048		Chicken nov gene 4th	1.08e-13
22	52	2.5	632	6	Q36034		Chicken nov gene frag	1.09e-12
23	44	2.1	204	1	N81164		Base substituted E.co	9.28e-09
24	44	2.1	91	9	Q51746	c	Oligonucleotide probe	9.28e-09
25	43	2.1	91	9	Q51746	c	Oligonucleotide probe	9.28e-09
26	43	2.1	204	1	N81164		Base substituted E.co	2.80e-08
27	39	1.9	56	5	Q33947		Microsatellite sequen	2.17e-06
28	38	1.8	60	5	Q33792		Microsatellite sequen	6.31e-06
29	36	1.7	63	5	Q33701		Microsatellite sequen	5.20e-05
30	35	1.7	1521	5	Q29724		Sequence of the first	1.47e-04
31	34	1.6	258	5	Q32866	c	Probe VAW409 for diag	4.12e-04
32	34	1.6	40	4	Q34094		Sequence of a microsa	4.12e-04
33	34	1.6	53	5	Q33627		Microsatellite sequen	4.12e-04
34	33	1.6	3871	2	N71302		HSV-1 gB and surround	1.14e-03
35	33	1.6	266	4	Q27097		Fragile X diagnostic	1.14e-03
36	33	1.6	840	4	Q29612		Encodes carp TLAIIbet	1.14e-03
37	33	1.6	49	4	Q34122		Sequence of a microsa	1.14e-03
38	33	1.6	59	4	Q34173		Sequence of a microsa	1.14e-03
39	32	1.5	327	10	Q77289		Human genome fragment	3.13e-03
40	32	1.5	50	4	Q34167		Sequence of a microsa	3.13e-03
41	32	1.5	870	6	Q39248		Sequence of cosmid cl	3.13e-03
42	32	1.5	62	4	Q34053		Microsatellite sequen	3.13e-03
43	31	1.5	79	4	Q33502		Sequence of microsate	8.49e-03
44	31	1.5	260	10	Q77444	c	Human genome fragment	8.49e-03
45	30	1.4	201	2	N70195	c	Streptomyces protease	2.27e-02

                        ALIGNMENTS

```
RESULT   1
ID   Q26422 standard; DNA; 2330 BP.
AC   Q26422;
DT   18-JAN-1993 (first entry)
DE   Gene for beta-IG-M2.
KW   Transforming growth factor beta; induced; CEF-10; v-src; chicken;
KW   embryo; fibroblasts; TGF-beta; ss.
OS   Mus musculus.
FH   Key             Location/Qualifiers
FT   CDS             204..1247
FT   /*tag= a
PN   EP-495674-A.
PD   22-JUL-1992.
PF   17-JAN-1992; 300429.
PF   18-JAN-1991; US-642991.
PR   10-JAN-1992; US-816270.
PA   (BRIM ) BRISTOL-MYERS SQUIBB CO.
PI   Brunner AM, Chinn J, Neubauer MG, Purchio AF;
DR   WPI; 92-243508/30.
```

DR   P-PSDB; R25565.
PT   TGF-beta induced gene family - encodes proteins involved in
PT   growth and differentiation effects of TGF-beta-1
PS   Claim 7; Fig 2; 35pp; English.
CC   The DNA encoding mouse beta-IG-M2 was obtd. from AKR-2B mouse cells
CC   induced with TGF-beta1 and cyclohexamide. Poly RNA extracted from
CC   these cells was used to create a cDNA library which was screened
CC   using two probes. The probes were prepd. from untreated AKR-2B mRNA
CC   and AKR-2B mRNA from cells treated with cyclohexamide and TGF-beta1.
CC   Hybridising colonies were isolated and two clones (beta-IG-M1 and
CC   beta-IG-M2) were then sequenced. The DNA encodes proteins that
CC   have a 80 and 50 percent homology respectively with the CEF-10
CC   protein induced by v-src in chicken embryo fibroblasts. The
CC   proteins encoded by TGF-beta induced genes are likely to be
CC   involved in mediation of the biological effects of TGF-beta
CC   relating to cell growth and differentiation.
CC   See also Q26421.
SQ   Sequence   2330 BP;   589 A;   598 C;   595 G;   548 T;

DB   4;   Score   921;   Match 80.8%;   QryMatch 44.4%;   Pred. No. 0.00e+00;
Matches 1544;   Conservative   0;   Mismatches 317;   Indels   51;   Gaps 29;

```
Db   241 tggtgctcctcgccctctgcacccggcctgctacgggccaggactgcagcgcgcaatgtc 300
         ||||| |||||||||||| ||||| ||  |||||| |||||  ||  ||  ||  || ||
Qy   170 TGGTCCTCCTCGCCCTCTGCAGCCGGCCGGCGCGGCGACTCGCGCGGCCCGGCCCGTGCC 229

Db   301 agtgcgcagccgaagcagcgccgcactgccccgccgccgcgggctgagcctggtgctggacggct 360
         |||| | ||| |||||||||||||| |||||||| ||  ||  |||||||||||||
Qy   230 GGTGCCCGGACGAGCCCGCGCCGCGTCCCGGCGGGCCGTGAGCCTCGTGCTCGGACGGCT 289

Db   361 gcggctgctgccgcgtctgcgcaagcagcagtgtgacggaagactgtgtacggagctggtgcg 420
         |||||||||||| || |||||||||||||||  ||| |||||| || ||||||||
Qy   290 GCGCGTGCTGCCCGGTCGCTGGGCCAAGCAGCAGTGGGCCGAGCTGTGCACCGCCGCACCCCT 349

Db   421 gcgaccacaaggcgcctctctcgcgatttcggctccccgccaaccgcaagattggag 480
         ||||| |||||||||||| || |||||| ||||||||||| || ||||||||| ||
Qy   350 GCGACCCGCACAAGGGCCTCTTCGTCGACTTCGGTCCGGCTCCGCCCACCGCCAACGATCGGCG 409

Db   481 tgtgcactgcaaagatggtgcacccgtgtcttcggtgggtcggtgtaccgcagcgtg 540
         | |||||||| |||| |||||| |||| |||||||||| ||||  ||||||||||| |
Qy   410 TGTGCACCGGCCAAAGATGGTCGCTCCCTGCACTTCGGTGCGGGTACGGGTGTACCCAGCGGAG 469

Db   541 agtccttcaaagcagctgcaaataccaatgcacttgccctggatgggccgtgggctgcg 600
         ||||||| ||||||| ||| ||||| |||  || |||  ||||||||||||| ||||||
Qy   470 AGTCCTTCCAGAGCAGCTGCAAGTGCAAGTGCCAGTGCCACGTGCCTGCCGCACCTGGGGCTGCA 529

Db   601 tgccccatgcagcatggacgtgcgcctaccacctgcccagccctcccgagaaggg 660
         || |||||| |||||||| ||| ||||  ||||| ||||||||| ||||||||
Qy   530 TGCCCCGTGCCACCATGGACGTTCGTCCTGCCCAGCCCTGACTGCCCAGGAGGAGGG 589

Db   661 tcaagctgctgggaaatgctgcgagagtggggtgtgtgacgagccaagaccgcacag 720
         ||||||| ||||| || | |||||||||||  | |||||||| ||||| ||  ||
Qy   590 TCAAGCTGCCCCGGGAATGCTGCTGCGAGGAGAGTGGTGGTGTGACAGCGGAACCCAAACCG 649

Db   721 cagttggcctgccctagctgcctaccgactggaagacacatttggccagaccaacta 780
         | ||||| ||||| || |||| ||  ||||||| |||||| | || ||||||||| ||
Qy   650 TGGTTGGGCCTGCCCTGCGGCCTTACCGACCGTTGGCCCAGACACCCGAACACGTTTGGCCCAGACCCCAACTA 709

Db   781 tgatgcagccaactgcctggtccgacaccacagagtggagcgctgttctaagacctgtg 840
         |||| ||||||||| || ||||||||| |||||||||||||||||||| |||||||| |
Qy   710 TGATTACAGCCAACTGCCTCGTCCGTCCAGACCCACAGAGTGGAGCGGCGCCTGCTGTTCCAGACCTGTC 769

Db   841 gaatgggcatctcacccagttaccaatgacaatacctctgcaagcaggaagcaga 900
         |||| |||||| |||||||||  || ||| |||   ||||||  | ||||  ||
```

```
Qy   770 GGATGGGCATCTCACCCGGTTACCAATGACAACGGCCTCCTCCAGGCTAGCAGAAGCAGA 829

Db   901 gcgcgcctgccgatggtcaggccctgcagaagctgcgaagctgacctgaggagcaacattaagaagggca 960
         ||||||| |||||||||| |||  ||||||| |||  ||||| | ||||||||||||
Qy   830 GCCGCCTCTGCATGGTCAGGCCTTGCCGAAGCTGACGACTGGAAGACAACATTAAGAAGGCCA 889

Db   961 aaaagtgcatccggacacctaaaatcgccaagcctgtcaagtttgagctttcggctgca 1020
         ||||||| || |  ||||| |||||||||||||||| |||||| ||| || |||| ||||
Qy   890 AAAAGTGCATCCGTACTCCCAAAATCTCCAAGCCTATCAAGTTTGAGCTTTCGGCTTCGGCTGCA 949

Db   1021 ccagtgaagacatacagggtcaagttctgcgggggtgtgcacagacggcgctgctgca 1080
         |||  ||||||| || |  |  ||||  || |||||| ||| ||||||||||||||||
Qy   950 CCAGCATGCAAGACATACCGCAGCTAAATTCGTGGAGTATGTACCGACGCGGCCGATGCTGCA 1009

Db   1081 caccgcacagaaccaccactctgcacgtggagttcaaatgcccggatggtggagatcatga 1140
         | |||||||||||||||||| || |||| ||||||||| || |||||| |||||| |  ||
Qy   1010 CCCCCACAGAACCACCACCCTGCCGGTGAGTTCAAGTGCCCCTGACGCGCGCGAGGTCATGA 1069

Db   1141 aaagaaatatgatgttcatcaagacctgtgcctgccattacacctgtctgggggacaatg 1200
         | |||||||||||| |||| |||| || |||||||||||| |  ||||||||| |||| ||
Qy   1070 AGAAGAACATCATGTTCATCAAGACCTGTGCCTGCCCGCCTGCCCATTACAACTGTCCCGGAGACAATG 1129

Db   1201 acatctttggtcctgctacacaggaagatgtacgggcacatggctaaagccaggaag 1260
         ||| | |||| ||| || || ||||||||| ||||  ||||||||||||||||||| |
Qy   1130 ACATCTTTGAATCGTCGTACTACAGGAAGATGTACGGAGACATGGCATGGAGCCAGCAGAGAG 1189

Db   1261 taaggacagaactcattagactaactaacttgaactgagttgcatctcatttcttctgt 1320
         | | |||||| ||| |||| || | ||||||||||||||  | |||| |||||||||||||
Qy   1190 TGAGAGACATTAACTCATTAGCAAGCTTGGACTTGAACTGATTCATCTCATTT--TCCGT 1247

Db   1321 aaaaacaattacagtagcacat-taatttaaatctgttttttaactaccgtgggaggaa 1379
         |||| ||| || |||||| | |  |||||  ||| ||||| |  |||  |||||||||
Qy   1248 AAAAATGATTTCAGTGACACAAGTTATTTAAATCTGTTTTTCTAACTG--G-GGGAAAAG 1304

Db   1380 ctatccccaaagtgagaacgttatgtcatggccatacaagtagtctgtcaacctcaga 1439
         |||||||||| ||||||| |||| | ||||||| || |||  || |||||||| |||||
Qy   1305 AT-TCCCACCCAATTCAAAACATTGTGCCCATGTCAA-ACAAAATAGTCTATCTTCCCCAGA 1362

Db   1440 cactggttcgagacagttcacctgacagtgt--t-cattagcgcacagtgccagaa 1496
         |||||| | ||| | ||||| |||||| || |  | ||||| |||||||| ||||| |
Qy   1363 CACTGGTTTGAAGAATGTTGAAGAATGTTAAGACTTGACACAGTGGAACTACATTAGTACACAGCACCAGAA 1422

Db   1497 cgcacactgaggtagtctctgaacagtggggaatgtgcaggagaagaaagagacaggtac 1556
         |||||| ||||| ||| ||  |||||| ||||||||| ||| | |||||| ||||||
Qy   1423 TGTATATTAAGGTGTGGCTTTAGGACGCAGTGG-GAGGTACCGGGCCCCG---GTTAG-TAT 1477

Db   1557 tagctgaggttattttaaaagcagcagtgtgcctacctttttggagtgtaaccggggagggg 1616
         | || ||||| ||| |||||||| ||| |||| ||||  ||  ||||||| ||||||||| |
Qy   1478 CATCAGATGCACTCTTATACG-AGTAATATGCCTCATTTGGAGTGTTAAGTTGAAGTATCTGG 1536

Db   1617 aaattatgacgcttgcgcagacagacctg---cctctagcgagctgagca-tgtgtcct 1672
         | |||| |||| |||||| |||||| | ||   |  || | ||| | ||||  ||  | ||
Qy   1537 AAATTTAGCCTGCTCACTGACCTGCCCTGTAGCCCCAGTCACAGCTAGGATGTGCATTCT 1596

Db   1673 cca-ctag-atgaggctgagtccagcacatttggccagaccaacta 1726
         ||| || |  | ||| || ||| | ||||   ||  |||||||||||
Qy   1597 CCAGCCATCAAGAGACTCAAGTGTCCTTAGCTCAGACAACAGCAGACTCAGCTCTG 1656

Db   1727 accattctgattccagtgacacttgtcaggagtccattctgttttagactggacag 1786
         | |||| ||||| | | |||||||||||| |||||||| ||||| || | |||||||
Qy   1657 AC-ATTCTGATTCGAATGACACTGTTCAGGAATCGGAATCGGTCGTGTCGATTAGACTGGACAG 1715

Db   1787 cttgtgcaagtaagttgcctgtaacaagccagattttt----attgatatttgtaaata 1842
         |||||||||||||||||| |||||||||||||||||||     |||||||||||||
```

PT   evaluate development and/or differentiation of tumours
PS   Claim 1; Fig 1; 67pp; French.
CC   The chicken nov gene is stimulated in avian nephroblastoma induced
CC   by avian myeloblastoma virus but not in normal adult kidney. The
CC   1975bp cDNA sequence was isolated from a gene bank prepared from
CC   chicken embryonic fibroblasts screened with a tumour-derived probe.
CC   The cDNA was cloned in Bluescript KS+ to form plasmid pClK which was
CC   labelled and used to screen a human placental DNA library. A nov-
CC   analogous gene was isolated from the human library. Nucleotide
CC   sequences which hybridise to the chicken nov gene under stringent
CC   conditions (i.e. 50% formamide, 5 x SCC) are claimed.
SQ   Sequence 1975 BP;   522 A;   451 C;   486 G;   516 T;

DB 6;   Score   264; Match 66.0%; QryMatch 12.7%; Pred. No. 2.19e-132;
Matches 575; Conservative   0; Mismatches 293; Indels 3; Gaps 1;

Db 151 cgccgcgctgcgcccgggagtgccgccgtgctgacgactgcggctgctgctggtgt 210
          ||||||||||| | |||||||||||||||||||||||||||||||||||||||| ||
Qy 248 CGCCGGCGCTGCCCGGCGGGGTGCGTGCGGTGACGCCTGCGTGCGGCTGCCTGCCGCGTCT 307

Db 211 gcgccgcgagcgcggggagagggctgctcccctctgctccctgcgacgagaggcggcggcc 270
          |||| ||||| |||| ||| || | | | ||||||||||||| || ||||
Qy 308 GCGGCAAGCAGCTGGGCGAGCTGTGCACCGAGCGGCGACCCTGCCACCCGCCACAAGGCC 367

Db 271 tctactcgacccgacgcgcccgagacggcccgagcggcgcgcggacctcgatcgatcgtgctgctggaag 330
          ||| ||| ||||||||||| |||| ||||||| | | || |||| | |!! |||| | !!!
Qy 368 TCTTCTGTGACTTCGGCTCCGCCCGACGCGCCAACCGCAGATCGGCGTGCGACCCGCCAAAGATG 427

Db 331 gggacaactgcgtgttcgatgggatgattaccgcaacgggagacgttccagcccagct 390
          | | ||| |||| ||| || ||||||||| | |||| ||||||| ! ||||| |||
Qy 428 GTGCTCCCTGCATCTTCGGTGTGCACGGTGTACCGCCAGCGGACGTCCTTCCAGACCAGCT 487

Db 391 gcaagtaccagtgcacctgccggacgggcagatcggtcggtcgcctgcctgccccgctgcaacctgg 450
          ||||||||||||| |||| |||| | |||||| | ! || |||| | | ||| | !!!
Qy 488 GCAAGTACCAGTGCACGTGCCTGACGGGCGGGGTGGGCTGGCCTGCCTGCGTGCGCCCCTGTCGCGCATGG 547

Db 451 gctgtgctccccggccaccgacctgccccttccgcgggaagatcgaagtcccggagagt 510
          | | || ||||||||||||| | |||| |||| |||| ! ! ! ||||| | |||
Qy 548 ACGTTCGTCGCTCGCCCAGCCTGACTCCCCTTCCCCAGGAGGGTCAAGCTGCCCGGGAAAT 607

Db 511 gctcgagaagtgggtgtcgacccccaggagatgaagtgctcctggggagtcttgctatgg 570
          ||||||| |||||| | || ! | | |! || ! !!!! | || | | ! !!!
Qy 608 GCTGCGAGAGTGGGTGTGTGACGAGCCCAAGGACCCAAACCGTGGTTGGGCCTGCCCCTCG 667

Db 571 ctgcataacagacaggaggccacacttgggatagacgtgtctgatcaattgta 630
          | || | || || || ||| || | ! !| ||||||| | |!!! !|
Qy 668 CGGCTTACCGACTGGAACACAGTTTGGCCCCAGACCCAACTATGATTAGAGCCAACTGCC 727

Db 631 ttgaacagacaacagaatggagtgcctgttccaaaagctggaatgggctttctaccc 690
          || | |||| ! | !| | ||| | |||| | | ||||| | | !!| ||| | || !||
Qy 728 TCGTCCAGACCACCAGTGGACGCGCGTCGTTCCAGACCTCGTGGGATGGGCATCTCCACCC 787

Db 691 gtgttaccaacagacaaatcagcagtgtgagatggtgagatggtggagtgcaggcaacacgacgttgcatgatga 750
          |||||||| ! !!! || ||||| | | || ||||| || ! ! !| |! | !|||
Qy 788 GGGTTACCCATCACAACCCCTCTCAGGCTAGACGACTAGACGACGAGCAGCAGCGCCCGTGTCCATGGTCA 847

Db 751 gacttgtgaaacgagaggagccatcctga---taagaaggaaaaaaatgtatccaaacaa 807
          | || !||| |! | |!!||| || ||| | !!!|||| | !|||||| !||
Qy 848 GGCCTTGCCGAAGCTCACCTGAAGAGAACATTAAGAGGCAAAAAGTGCATCCTACTC 907

Db 808 agaatccatgaaagctgttcgtttgaatacaagaagctgaccagtgtgcagacttaca 867
          ||| || | ! !!|| ||| !!||| ||| !! | |!||||| | |!!! !|
Qy 908 CCAAAATCCAAGCCTATCAAGTTTGAGCTTTGGCTGCGACCCATGAAGAGACATACC 967

Qy 1716 CTTGTGGCAAGTGAATTTGCCTGTAACAAGCCAGATTTTTTAAAATTTATATTGTAAATA 1775

Db 1843 ttgtg-----ga-tatatatatatatatatatttgtacagttatctaagttaattt 1896
           ||||| | | | | ||||||||||| ||||||||||||||
Qy 1776 TTGTGTGTGTGTGTGTGTATATATATATATATATGTACAGTTATCTAAGTTAATT 1835

Db 1897 aaagtcatttgtt--tttgt-tttaagtgcttt-tgg---gattttaaactgatagcctc 1949
           ||||| ||||| ||| | | |||| ! !|| || ! | ||||||
Qy 1836 AAAGTGTTCGCCTTTTATTTTTGTTTTTAATGCTTTGATATTTCAATGTTAGCCTC 1895

Db 1950 aaatccaaacaccataggtaggacacgaagcttatctg-tgattcaaaaca--aaggag 2006
           || |||||||||||||||||| ||||| | !| |||| | !|||| || | |!
Qy 1896 AATTTCTGAACACCATAGGTAGAATGTAAAGCTTGTCTGATCGTTCAAAGCCATGAAATGG 1955

Db 2007 atactgcagtgggaattgtgacctgagtgactctctgtcagaacaacaaatgctgtgca 2066
           |||| !| !|||| ||| | | |!| | !| | |||| |!|!! !| !||||
Qy 1956 ATACTTATATGGAAATTCTGCTCAGATAGAATGACAGCTCCCTCAAAACAGATTGTTTGCA 2015

Db 2067 --ggtgataaagctatgtattggaagtcagattcctag-taggaaatgtggt 2115
           |! || !||||| || | | | |||||||||| |||||||||||||
Qy 2016 AAGGGGAGGCCATCAGTCGTCTGCCAGGCTGATTTCTGAGGCTAGGAAATGTGGT 2067

RESULT 2
ID   Q36031 standard; cDNA; 1975 BP.
AC   Q36031;
DT   24-MAY-1993 (first entry)
DE   Chicken nov coding sequence.
KW   avian nephroblastoma; avian myeloblastoma virus; ss.
OS   Gallus domesticus.
FH   Key   Location/Qualifiers
FT   CDS   24..1079
FT   /*tag= a
FT   polyA_signal   1814..1818
FT   /*tag= b
FT   polyA_signal   1932..1937
FT   /*tag= c
FT   sig_peptide   24..95
FT   /*tag= d
FT   exon   1..92
FT   /*tag= e
FT   /number= 1
FT   exon   93..321
FT   /*tag= f
FT   /number= 2
FT   exon   322..573
FT   /*tag= g
FT   /number= 3
FT   exon   574..782
FT   /*tag= h
FT   /number= 4
FT   exon   783..1975
FT   /*tag= i
FT   /number= 5
PN   W09300430-A.
PD   07-JAN-1993.
PF   25-JUN-1992; F00589.
PR   25-JUN-1991; FR-007807.
PA   (CNRS) CENT NAT RECH SCI.
PI   Martinerie C, Perbal B;
WPI; 93-036377/04.
DR   P-PSDB; R31599.
PT   Nucleotide sequences hybridising to regions of chicken nov gene -
PT   useful as probes for detecting complementary sequences to

```
Db    868 acccctcgttactgtggcctctgccaatgatgggcgatgctgctgtacccacacaacaccaaaa 927
          | ||  |  |  |||||| |||  ||  ||  |||||||| |||||  ||||| |||| |||| |
Qy    968 GAGCTAAATTCTGTGGAGTATGTACCGACGGCCGATGCTGCACCCCCCACAGAACCACCA 1027

Db    928 cgattcaagttgagttccgctgcctcaggcaaattcctaaaaaagccaatgatgttga 987
          | | ||  || ||||| || ||| || ||  ||| |  || | |||| ||| |||||||| |
Qy   1028 CCCTGCCGGTGGAGTTCAAGTGCCCTCACGCGCCAGCGTCATGAAGAAGAACATGATGTTCA 1087

Db    988 tcaataacctgtgtctgtcatggtaactgtcc 1018
          |||| |||||| |  || |  |||||||||
Qy   1088 TCAAGACCTGTGTGCCTGCCATTACAACTGTCC 1118
```

RESULT   3
ID   Q36046 standard; cDNA; 619 BP.
AC   Q36046;
DT   24-MAY-1993 (first entry)
DE   Human sequence XXVII homologous to chicken nov gene.
KW   avian nephroblastoma; avian myeloblastoma virus;
KW   stringent hybridisation; ss.
OS   Homo sapiens.
PN   WO9300430-A.
PD   07-JAN-1993.
PF   25-JUN-1992; F00589.
PR   25-JUN-1991; FR-007807.
PA   (CNRS) CENT NAT RECH SCI.
PI   Martinerie C, Perbal B;
DR   WPI; 93-036377/04.
PT   Nucleotide sequences hybridising to regions of chicken nov gene –
PT   useful as probes for detecting complementary sequences to
PT   evaluate development and/or differentiation of tumours
PS   Disclosure and Claim 26; Page 41; 67pp; French.
CC   In the claims, sequence XXVII is described as part of the chicken
CC   nov gene exon 4. In the disclosure (which contains more reliable
CC   descriptions of other sequences mentioned in the claims), sequence
CC   XXVII is described as a preferred nucleotide sequence of the
CC   invention which is found in humans and can hybridise to a portion
CC   of the chicken nov gene 3rd. exon under stringent conditions (i.e.
CC   50% formamide, 5 x SCC).
SQ   Sequence 619 BP;   114 A;   197 C;   175 G;   133 T;

DB 6; Score   240; Match 98.8%; QryMatch 11.6%; Pred. No. 2.44e-118;
Matches   243; Conservative   0; Mismatches   3; Indels   0; Gaps   0;

```
Db    99 gccaaagatgtgctccctgatcttcggtggtacggtgtaccgcagccgagagtccttc 158
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   418 GCCAAAGATGGTGCTCCCTGATCTTCGGTGGTACCGGTGTACCGCAGCCGAGAGTCCTTC 477

Db   159 cagagcagctgcaagtaccactgcagctgctgacggcgggacgggcggtggctgcatgccccctg 218
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   478 CAGAGCAGCTGCAAGTACCAGTGCACGTGCAGCGTGCCTGGACGTGGGCGGTGGCTGCATGGCCCCTG 537

Db   219 tgcagcatgacgttcgtctgccagccctgactgccccttcccgaggagggtcaagctg 278
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   538 TGCAGCATGACGTTCGTCTGCCAGCCCTGACTGCCCCTTCCCGGAGAGGGTCAAGCTG 597

Db   279 cccgggaaatgctgcgaggagtgggtgtgtgacgagcccaaaccgtccttggg 338
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   598 CCCGGGAAATCTCGCGAGGACGTGGGTGTGTGACGAGCCCAAGGACCCAAACCGTGTTGGG 657

Db   339 cctgcc 344
         ||||||
Qy   658 CCTGCC 663
```

RESULT   4
ID   Q36032 standard; cDNA; 684 BP.
AC   Q36032;
DT   24-MAY-1993 (first entry)
DE   Chicken nov gene fragment II.
KW   avian nephroblastoma; avian myeloblastoma virus;
KW   stringent hybridisation; ss.
OS   Gallus domesticus.
PN   WO9300430-A.
PD   07-JAN-1993.
PF   25-JUN-1992; F00589.
PR   25-JUN-1991; FR-007807.
PA   (CNRS) CENT NAT RECH SCI.
PI   Martinerie C, Perbal B;
DR   WPI; 93-036377/04.
PT   Nucleotide sequences hybridising to regions of chicken nov gene –
PT   useful as probes for detecting complementary sequences to
PT   evaluate development and/or differentiation of tumours
PS   Claim 1; Page 26; 67pp; French.
CC   The chicken nov gene is stimulated in avian nephroblastoma induced
CC   by avian myeloblastoma virus but not in normal adult kidney. A
CC   1975bp cDNA sequence was isolated from a gene bank prepared from
CC   chicken embryonic fibroblasts screened with a tumour-derived probe.
CC   Fragment II is derived from the nov gene; nucleotide sequences
CC   which hybridise to Fragment II under stringent conditions (i.e.
CC   50% formamide, 5 x SCC) are claimed.
SQ   Sequence 684 BP;   141 A;   197 C;   226 G;   120 T;

DB 6; Score   215; Match 67.6%; QryMatch 10.4%; Pred. No. 9.12e-104;
Matches   413; Conservative   0; Mismatches 198; Indels   0; Gaps   0;

```
Db    26 cgccggcgctgcgccccgggagtgccgccgctgcggacggctgccgctgctgcctggtgt 85
         |||||||||||| |  | |||  ||  |||| |||||| |||||| ||||||||| ||||
Qy   248 CGCCGGCGCTGCCCGGCGGGCGGCGTGAGCCTGTCCTGCGACGGCCTGCGCTGCTGCCCGCGTCT 307

Db    86 gcgcccgcagccgacgcggcgagagctgctccccctctgctgcccctgcgacgagacggcggcc 145
         ||||| ||||||||| ||||| |||  |  || |  ||||||||| | ||||||||
Qy   308 GCCCCAAGCAGCTGGGCGAGCTGTGTCCACCGGCACGCCGCGCACCCCTGCGACCCGCACAAGGGCC 367

Db   146 tctactgcgaccgcggccccgaggacagacggcgcgcgccggcatctgcatggtgctggaag 205
         |||  | || || |||||  ||  |  ||  | |   ||||| |||  || | |  ||| ||
Qy   368 TCTTCTGTGACTTCGGCTCTCCCCGGCCAACCGCAGCATCGGCCTGTGCACCCGCCAAAGATG 427

Db   206 gggacaactgcgtgttcgatgggatgatttaccgcaacgggagacgttcagcccagct 265
         |||  |  || ||| || |  |||  |  |  |||| ||| ||  | ||||| | |||
Qy   428 GTGCTCCCTGCCATCTTCGGTGCTGTACGGTGTACCGCAGCCGGAGAGTCCTTCCAGACCAGCT 487

Db   266 gcaagtaccagtgcacctgccggacagggcacggcagatcggtgctgcccccgctgcaacctgg 325
         |||||||||||||||| || |  |  || ||||  | ||| | ||||| |  |  ||| ||| ||
Qy   488 GCAAGTACCAGTGCACGTGCCTGACGGGGCCGGTGGCTGCTGCATGCCCCCTGTGCAGCATGG 547

Db   326 gctgctgctcccggcccccgactgccccttcccgcggaagatcgaagtcccggagagt 385
         |||  |  | | |||||||||||||||||||||| |||  |  ||  | |||||| |||
Qy   548 AGTTCGTCGCCCAGCCCTGACTGCCCCTTCCCGGAGAGGGTCAAGCTGCCCGGGAAAT 607

Db   386 gctggagaagtgggtgtgcgacccccaggatgaagtgctcctggggaggcttgctatgg 445
         |||| |  | |||||||| ||| || || |  || ||| |   |   |||  ||| |||
Qy   608 GTGCTCCCTGCCATCTTCGGTGCTGTACGGTGTACCGCAGCCCAAGGACCCAAACCGTGTTGGG 667

Db   446 ctgcatacagacaggagccacacttggatagacgtgtctgattcaagtgccaattgta 505
         ||| |  |  || | || |  ||  |  |  |  || | | | |  ||| |  |  ||
Qy   668 CGGCTTACCGACTGAAGACACGTTTGCCCCAGACCCAACTATGATTAGAGCCAACTGCC 727
```

```
Db   506 ttgaaacagcaacagaatggagtgcttgttccaaaagcgtgtggaatgggctttctaccc 565
Qy   728 TGGTCCAGACCACGAGTGCAGCCGCCTCGTTCCAAGACCTCTGGGATGGGCATCTCCACCC 787

Db   566 gtgttaccaacagaaaatcagcagtcgtgagatggtgaagcagacacgacttgcatgatga 625
Qy   788 GGGTTACCAATGACAACGCCCTCCTGCAGGCTAGAGAAGCAGCAGCCCGCCTGTCGCATGGTCA 847

Db   626 gacttgtgaa 636
Qy   848 GGCCTTCGGAA 858
```

RESULT   5
ID   Q36044 standard; cDNA; 461 BP.
AC   Q36044;
DT   24-MAY-1993 (first entry)
DE   Chicken nov gene exon 3-4 fragment XXII.
KW   avian nephroblastoma; avian myeloblastoma virus;
KW   stringent hybridisation; ss.
OS   Gallus domesticus.
PN   WO9300430-A.
PD   07-JAN-1993.
PF   25-JUN-1992; F00589.
PR   25-JUN-1991; FR-007807.
PA   (CNRS) CENT NAT RECH SCI.
PI   Martinerie C, Perbal B;
DR   WPI; 93-036377/04.
PT   Nucleotide sequences hybridising to regions of chicken nov gene –
PT   useful as probes for detecting complementary sequences to
CC   evaluate development and/or differentiation of tumours
PS   Claim 20; Page 38; 67pp; French.
CC   The chicken nov gene is stimulated in avian nephroblastoma induced
CC   by avian myeloblastoma virus but not in normal adult kidney. A
CC   1975bp cDNA sequence (Q36031) was isolated from a gene bank prepared
CC   from chicken embryonic fibroblasts screened with a tumour-derived
CC   probe. Fragment XXII is part of the 3rd and 4th exons of the nov gene;
CC   nucleotide sequences which hybridise to Fragment XXII under stringent
CC   conditions (i.e. 50% formamide, 5 x SCC) are claimed. The claimed
CC   sequences preferably encode a protein with the sequence XXIII
CC   (R31608) or encode a protein with at least 60% homology to it.
SQ   Sequence 461 BP; 109 A; 116 C; 143 G;   93 T;

DB 6; Score 160; Match 68.9%; QryMatch 7.7%; Pred. No. 5.23e-72; Indels 0; Gaps 0;
Matches 292; Conservative 0; Mismatches 132;

```
Db    17 ctgcgtgttcgatggatgatttaccgcaacgggagacgttccagccagctgcaagta 76
Qy   435 CTGCATCTTCGGTGGTACGGTGTACCGCACCGGAGAGTCCTTCCAGAGCAGCTGCAAGTA 494

Db    77 ccagtgcacctgccggacgggacagatccggtgcctgcccccgctgcaacctgggcctgct 136
Qy   495 CCAGTGCACCTGCCTGACGGCGGCGGTGGGCTGCATGCCCCTGTGCAGCATGACCGTTCG 554

Db   137 gctcccggccccgactgccccttcccgcggaaatcgaagtccccggagagttgctgcga 196
Qy   555 TCTGCCCAGCCCTGACTGCCCCTTCCCGAGGAGGTCAAGCTGCCCGGGAAATGCTGCGA 614

Db   197 gaagtgggtgcgaccccaggatgaagtgctcctggaggctttgctatggctgcata 256
Qy   615 GGAGTGGGTGTGTGACCAGGACCAAACCGTGGTTGGGCCTGCCCTCGCGGCTTA 674

Db   257 cagacaggaggccacacttggatagacgtgtctgattcaagtgcaattgattgaaca 316
```

```
Qy   675 CCGACTGGGAAGACACGTTCGCCCAGACCCAACTATGATTAGAGCCAACTGCCTGTCCA 734
Db   317 gacaacagaatggagtgcttgttccaaaagctgtggaatggctttttctacccgtttac 376
Qy   735 GACCACAGAGTGGAGCCGCCTGTTCCAAGACCTGTGGGATGGCATCTCCACCCGGGTTAC 794

Db   377 caacagaaatcagcagtgtgagcagacagcacgacttgcatgatgagacacctg 436
Qy   795 CAATGACAACGCCCTCCTGCAGGCTAGAGAAGCAGCAGCCCGCCTGTCGCATGGTCAGGCCCTTG 854

Db   437 tgaa 440
Qy   855 CGAA 858
```

RESULT   6
ID   Q36035 standard; cDNA; 252 BP.
AC   Q36035;
DT   24-MAY-1993 (first entry)
DE   Chicken nov gene exon 3 fragment VIII.
KW   avian nephroblastoma; avian myeloblastoma virus;
KW   stringent hybridisation; ss.
OS   Gallus domesticus.
PN   WO9300430-A.
PD   07-JAN-1993.
PF   25-JUN-1992; F00589.
PR   25-JUN-1991; FR-007807.
PA   (CNRS) CENT NAT RECH SCI.
PI   Martinerie C, Perbal B;
DR   WPI; 93-036377/04.
DR   P-PSDB; R31602.
PT   Nucleotide sequences hybridising to regions of chicken nov gene –
PT   useful as probes for detecting complementary sequences to
PT   evaluate development and/or differentiation of tumours
PS   Claim 7; Page 30; 67pp; French.
CC   The chicken nov gene is stimulated in avian nephroblastoma induced
CC   by avian myeloblastoma virus but not in normal adult kidney. A
CC   1975bp cDNA sequence (Q36031) was isolated from a gene bank prepared
CC   from chicken embryonic fibroblasts screened with a tumour-derived
CC   probe. Fragment VIII is derived from the 3rd. exon of the nov gene;
CC   nucleotide sequences which hybridise to Fragment VIII under stringent
CC   conditions (i.e. 50% formamide, 5 x SCC) are claimed. The claimed
CC   sequences preferably encode a protein with at least 70% homology to
CC   R31602 which is encoded by Fragment VIII.
SQ   Sequence 252 BP;   43 A;   76 C;   88 G;   45 T;

DB 6; Score 106; Match 77.0%; QryMatch 5.1%; Pred. No. 1.64e-41; Indels 0; Gaps 0;
Matches 151; Conservative 0; Mismatches 45;

```
Db    18 ctgcgtgttcgatgatgatttaccgcaacgggagagacgttccagccagctgcaagta 77
Qy   435 CTGCATCTTCGGTGGTACGGTGTACCGCACCGGAGAGTCCTTCCAGAGCAGCTGCAAGTA 494

Db    78 ccagtgcacctgccggacgggacagatcggtgcctgcccccgctgcaacctgggcctgct 137
Qy   495 CCAGTGCACCTGCCTGACGGCGGCGGTGGGCTGCATGCCCCTGTGCAGCATGACCGTTCG 554

Db   138 gctcccggccccgactgcccctttcccgcggaaatcgaagtccccggagagtgctgcga 197
Qy   555 TCTGCCCAGCCCTGACTGCCCCTTCCCGAGGAGGTCAAGCTGCCCGGGAAATGCTGCGA 614

Db   198 gaagtgggtgtgcgac 213
Qy   615 GGAGTGGGTGTGTGAC 630
```

RESULT 7
ID   Q36045 standard; cDNA; 196 BP.
AC   Q36045;
DT   24-MAY-1993 (first entry)
DE   Chicken nov gene exon 3 fragment XXIV.
KW   avian nephroblastoma; avian myeloblastoma virus;
KW   stringent hybridisation; ss.
OS   Gallus domesticus.
PN   WO9300430-A.
PD   07-JAN-1993.
PF   25-JUN-1992; F00589.
PR   25-JUN-1991; FR-007807.
PA   (CNRS) CENT NAT RECH SCI.
PI   Martinerie C, Perbal B;
DR   WPI; 93-036377/04.
PT   Nucleotide sequences hybridising to regions of chicken nov gene -
PT   useful as probes for detecting complementary sequences to
PT   evaluate development and/or differentiation of tumours
PS   Claim 24; Page 40; 67pp; French.
CC   The chicken nov gene is stimulated in avian nephroblastoma induced
CC   by avian myeloblastoma virus but not in normal adult kidney. A
CC   1975bp cDNA sequence (Q36031) was isolated from a gene bank prepared
CC   from chicken embryonic fibroblasts screened with a tumour-derived
CC   probe. Fragment XXIV (Q36045) is part of the 3rd exon of the nov gene;
CC   nucleotide sequences which hybridise to Fragment XXIV under stringent
CC   conditions (i.e. 50% formamide, 5 x SCC) are claimed. (N.B. the
CC   sequence XXIV is described as an amino acid sequence in the claims
CC   but is described correctly in the disclosure).
SQ   Sequence 196 BP; 32 A; 65 C; 66 G; 33 T;

DB   6; Score   106; Match 77.0%; QryMatch 5.1%; Pred. No. 1.64e-41;
Matches 151; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

Db   1 ctcgtgttcgatggatgattaccgcaacgggagacgttccagcccagctgcaagta 60
         ||||| ||||| ||| | | | |||||| ||| ||||| ||||||||||||||
Qy 435 CTCGATCTTCGGTGTACGGTGTACCCGACCGGAAGAGCTCCTTCCAGAGCAGCTGCAAGTA 494

Db   61 ccagtgcacctgccgggacgggacagcagatcgggtgcctgccccgtgcaacctgggcctgct 120
         ||||||| |||||||| ||||||| ||| | |||||| ||| | | | | |
Qy 495 CCAGTGCACGTGCCTGACGCGGCGGCGGTGGCCTGGCTGCGATCCCCTGTGCAGCATGGACGTTCG 554

Db 121 gtccccgccccgactgccccttccgcggaagatcgaagtccccggagagtgctgcga 180
         | ||| |||| |||||||||||| ||| | | | | | ||||||| | ||||||
Qy 555 TCTGCCCCAGCCCTGACTGCCCCCTTCCCGAGGAGGGTCAACCTGCCCCGGAAATGCTGCGA 614

Db 181 gaagtgggtgtgcgac 196
         | |||||||||| |||
Qy 615 GGAGTGGGTGTGTGAC 630

RESULT 8
ID   Q36049 standard; DNA; 131 BP.
AC   Q36049;
DT   24-MAY-1993 (first entry)
DE   Fragment XXXII encodes protein homologous to chicken nov protein.
KW   avian nephroblastoma; avian myeloblastoma virus;
KW   stringent hybridisation.
OS   Homo sapiens.
PN   WO9300430-A.
PD   07-JAN-1993.
PF   25-JUN-1992; F00589.
PR   25-JUN-1991; FR-007807.

PA   (CNRS) CENT NAT RECH SCI.
PI   Martinerie C, Perbal B;
DR   WPI; 93-036377/04.
PT   Nucleotide sequences hybridising to regions of chicken nov gene -
PT   useful as probes for detecting complementary sequences to
PT   evaluate development and/or differentiation of tumours
PS   Disclosure; Page 42; 67pp; French.
CC   The invention includes nucleotide sequences which encode amino acid
CC   sequences with at least 80% homology to sequence XXX (R31611), i.e
CC   to part of the chicken nov protein encoded by the 4th. exon. Such
CC   sequences preferably encode the amino acid sequence XXXI (R31612)
CC   and have the nucleotide sequence XXXII.
SQ   Sequence 131 BP; 30 A; 40 C; 35 G; 26 T;

DB   6; Score   103; Match 95.6%; QryMatch 5.0%; Pred. No. 7.54e-40;
Matches 129; Conservative 0; Mismatches 2; Indels 4; Gaps 2;

Db   1 ctgctctgcagcttaccgactggaagacacgtttggccagaccaactatgattagag 60
         |||| ||||| |||||||||||||||||||||||||||||||||||||||||||||||
Qy 659 CTGCCCTCGCGGCTTACCGACTGGAAGACACGTTTGCCCAGACCCAACTATGATTAGAG 718

Db   61 ccaac---ctggtccagac-acagagtggagcgcctgttccaagacctgtgggatgggca 116
         ||||| |||||||| ||||||||||||||||||||||||||||||||||||||
Qy 719 CCAACTGCCTGGTCCAGACCACCAGAGTTGGAGCGCCTGTTCCAGACCTGTGGGATGGGCA 778

Db 117 tctccacccggtta 131
         ||||||||||||||
Qy 779 TCTCCACCCGGTTA 793

RESULT 9
ID   Q26421 standard; cDNA; 2028 BP.
AC   Q26421;
DT   18-JAN-1993 (first entry)
DE   Gene for beta-IG-M1.
KW   Transforming growth factor beta; induced; CEF-10; v-src; chicken;
KW   embryo; fibroblasts; TGF-beta; ss.
OS   Mus musculus.
FH   Key   Location/Qualifiers
FT   CDS   186..1322
FT   /tag= a
PN   EP-495674-A.
PD   22-JUL-1992.
PF   17-JAN-1992; 300429.
PR   18-JAN-1991; US-642991.
PR   10-JAN-1992; US-816270.
PA   (BRIM) BRISTOL-MYERS SQUIBB CO.
PI   Brunner AM, Chinn J, Neubauer MG, Purchio AF;
DR   WPI; 92-243508/30.
DR   P-PSDB; R25565.
PT   TGF-beta induced gene family - encodes proteins involved in
PT   growth and differentiation effects of TGF-beta-1
PS   Claim 6; Fig 1; 35pp; English.
CC   The DNA encoding mouse beta-IG-M1 was obtd. from AKR-2B mouse cells
CC   induced with TGF-beta1 and cyclohexamide. Poly RNA extracted from
CC   these cells was used to create a cDNA library which was screened
CC   using two probes. The probes were prepd. from untreated AKR-2B mRNA
CC   and AKR-2B mRNA from cells treated with cyclohexamide and TGF-beta1.
CC   Hybridising colonies were isolated and two clones (beta-IG-M1 and
CC   beta-IG-M2) were then sequenced. The DNA encodes proteins that
CC   have a 80 and 50 percent homology respectively with the CEF-10
CC   protein induced by v-src in chicken embryo fibroblasts. The
CC   proteins encoded by TGF-beta induced genes are likely to be
CC   involved in mediation of the biological effects of TGF-beta

```
CC   relating to cell growth and differentiation.
CC   See also Q26422.
SQ   Sequence 2028 BP;   510 A;   535 C;   528 G;   455 T;

DB  4;  Score   88;  Match 61.1%;  QryMatch 4.2%;  Pred. No. 1.28e-31;
Matches  243;  Conservative   0;  Mismatches 155;  Indels   0;  Gaps   0;

Db  861 aaatgcatcgttcagacaccagtcttggtcccagtgctccaagagctgcggaactggcatc 920
Qy  721 AACTGCCTGGTCCAGACCACAGAGTGAGCCCCTGTTCCAAGACCTGTGGGATGGGCATC 780

Db  921 tccacacgagttaccaatgacaaccagagtgccggcctgtgaaagagaccggatcgt 980
Qy  781 TCCACCCGGTTACCAATGACAACGCCCTTCCTCAGGCTAGAGAAGCAGAGCCCCTGTGC 840

Db  981 gaagtgcgtccttgtggacaaccagtgtacagcagcctaaaaagggcaagaaatgcagc 1040
Qy  841 ATGTCAGGCCTTGCGAAGCTGACCTGGAAGAGAACATTAAGAAGGCAAAAAGTGCATC 900

Db 1041 aagacaagaaatcccagaaccagtcagattacttatgcaggatgctccagtgtcaag 1100
Qy  901 CGTACTCCAAAATCTCAAGCCTATCAAGTTTGAGCTTTCTGGCTCACCACCATGAAG 960

Db 1101 aaataccggcccaaaatactggcgctctgcgtagatggccggtgctgcacacctgcag 1160
Qy  961 ACATACCGAGCTAAATTCTGTGGATATGTACCGACGGCCGATGCTGCACCCCCACAGA 1020

Db 1161 accagaactgtgaagatgcggttccgatgcgaagatggagagatgtttccaagaatgtc 1220
Qy 1021 ACCACCACCCTGCCGGTGGAGTTCAAGTGCCCTGACGGCCAGGTCATGAAGAAGAACATG 1080

Db 1221 atgatgatccagtcctgcaaatgtaactacaactgccc 1258
Qy 1081 ATGTTCATCAAGACCTGCCTGCCATTACAACTGTCC 1118

RESULT  10
ID  Q10572 standard; DNA; 1047 BP.
AC  Q10572;
DT  09-APR-1991 (first entry)
DE  Human Natriuretic Peptide Receptor B.
KW  NPRB; ANP; BNP; CNP; kidney failure; heart failure; protein kinase;
KW  hyperaldosteronism; glaucoma; guanyl cyclase.
OS  Homo sapiens.
FH  Key        Location/Qualifiers
FT  Peptide     1..22
FT  /label= signal sequence
FT  Protein    12
FT  /label= mature NPBR
FT  Domain     23..455
FT  /label= extracellular domain
FT  /note= "binds natriuretic peptides A,B and C]"
FT  Domain     456..456
FT  /label= transmembrane domain
FT  Domain     479..1047
FT  /label= cytoplasmic domain
FT  /note= "GC and protien kinase activity"
FT  Modified -site 24..26
FT  /label= N-glycos_site
FT  Modified -site 35..37
FT  /label= N-glycos_site
FT  Modified -site 161..163
FT  /label= N-glycos_site
FT  Modified -site 195..197
```

```
FT   /label= N-glycos_site
FT   Modified -site 244..246
FT   /label= N-glycos_site
FT   Modified -site 277..279
FT   /label= N-glycos_site
FT   Modified -site 349..351
FT   /label= N-glycos_site
FT   Modified -site 600..602
FT   /label= N-glycos_site
PN   WO910292-A.
PD   10-JAN-1991.
PF   22-JUN-1990; U03586.
PR   23-JUN-1989; US-370673.
PA   (GETH ) GENENTECH INC.
PI   Chang M, Goeddel D, Lowe D;
DR   WPI; 91-036711/05.
DR   N-PSDB; Q10324.
PT   Natriuretic protein receptor B - for diagnosis and treatment of
PT   kidney failure, heart failure, hyperaldosteronism, glaucoma etc.
PS   Claim 3; Fig 1; 49pp; English.
CC   The sequence was derived from the DNA encoding natriuretic peptide
CC   receptor B, NPRB, having guanyl cyclase (GC) activity and protein
CC   kinase activity. The DNA can be inserted into expression vectors
CC   for the prodn. of the protein, opt. after being mutated to produce
CC   NPRB analogues. The protein has a mol wt. of 115 kD (calculated Mr=
CC   114,952).  The protein (or variants) can be used in treatment of
CC   natriuretic peptide disorders, and also to isolate peptides using
CC   affinity chromatography.  Antibodies with affinity for NPRB can
CC   also be prepd.
SQ   Sequence 1047 BP;   87 A;   15 C;   83 G;   51 T;

DB  2;  Score   80;  Match 9.5%;  QryMatch 3.9%;  Pred. No. 2.72e-27;
Matches   80;  Conservative 223;  Mismatches 532;  Indels   7;  Gaps   7;

Db  186 gvnnanngsnnsvnhnvyarnngnnnathnnrangrnvyncgnnmnhnnnnanrnnn 245
Qy  215 GCAGCGGCCCGCCCGGTGCCCGGACCAGCCCGG-CGCCGGCCGCCCGGGGGCCGGTGAGC 273

Db  246 tngdyvnnyndvngnsnragntratgrnwndnrtrnnananrnannntvnvntyrnnnnn 305
Qy  274 CTCGTGCTGGACGGGTGCGGCGTGCTGCCGCGGCTCGCCCAAGCAGCTGGGCGAGCTGTGC 333

Db  306 ynnnnnrnnrarndngvnngnsmnnnagcnydgnnnyanvnnntnnngtnndgnrnv 365
Qy  334 ACCGAGCCCGACCCCTGCGACCCCGCACAAGGGCCCTCTTCTGTGACTTCGGCTCCCCGGCC 393

Db  366 nkmngrryhgvtgnvvmdknndrntdhvnwamgdndsgdnnnaahysganknnwwtgrnn 425
Qy  394 AACCGCA-AGAT-CGGCGTGCTGCACCCGCCAAAGATGTGCTCCCTGCCTCCTGGTGGTA 451

Db  426 nwvkgannsdnnncandnddnscdktnnstnanvangtgntnmngvssnnnrknmnnk 485
Qy  452 CGGTGTACCCAGCCGGAGAGTCCTTCCAGACAGCCAGCTCCAACTACCACGTGCACCGTGCCTGG 511

Db  486 nnasmnwrn-rwnnnnngnsnryhkgagsrntnsnrgssygsnmtahgkynnnantghnk 544
Qy  512 ACGGGGCGTGGCTGGCTGCATGCCCGTGCAGCATGGACGTTGCTGTCGCCCAGCCCTGACT 571

Db  545 gnvvankhvnkkrnnntrnvnnnnkhmrdvnnnhnttrnngacndnnncnvtnycnrgsn 604
Qy  572 GCCCCTTCCCAGGAGGGTCAAGCTGCCCCGGAAATCGTCGCCAGGACTGGGTGTGTCACG 631

Db  605 ndnnnndsnnndwmryysnnndnvkgmannhnsnnsshgsnkssncvvdsrnvkntdyg 664
```

Qy   632 AGCCCAAGAGACCAAACCGTGGTTGGCCTGCCCTCCGCGGCTTACCGACTGGAAGACACGT 691

Db   665 nasnrstanndanayakknntannnnsgnnnnttgmnaadvysngnnnnnanrsgnn 724
          : : | : | |:|:: || :   : |

Qy   692 TTGGCCCAGACCCAACTATGATTAGAGCACCAACTGCCTGGTCCACCACAGAGTGGAGCC 751

Db   725 ynngndnsnknnvkvrngnrnynrnsndrtnnnnvnnvnmnrcwandnanrndngnnkg 784
          : | : : :: | : | | | | | | : | :||::|:

Qy   752 CCTGTTCCAAGACCTG-TGGGATGGGCATCTCCACCCGGGTTACCAATGACAACGCCTCC 810

Db   785 nnrrnnknggtsnndnnnrmnnyannnnknvnnrtnaynnnkrkanannynnnhsvan 844
          : | : :: ::   : : : :|:| : || : : : :

Qy   811 TCCAGGCTAGAGAAGCAGGACGCGC-CTGTCCATGGTCAGGCCTTGCCAAGCTGACCTGGA 869

Db   845 nnkrgntvnanandsvtnynsdnvgntansanstnmnvtnnndnytcndanndnndvyk 904
          : | : : : : | :|| :: | :|| :|| :| ::

Qy   870 AGAGAACATTAAGAGAGGCAAAAAGTCGATCCTACTCCCCAAAATCTCCAAGCCTATCAA 929

Db   905 vntngdaymvvsgnngngrnrhannnarmananandavsenrnrhrnhdnnrnrngvh-tg 963
          : | : : : :: | :| |:|:|: : : : : :: ::||

Qy   930 GTTTGAGCTTTCTGGCTGCCACCAGCATGAAGACATACCGAGCTAAATTCTGTCGAGTATG 989

Db   964 nvcagvvgnkmnrycnngdtvntasrmnsngnanknhvsttkdandnngcnnnnnrgdv 1023
          :| |:| : ::| :   :   : : ||::|: ::|:| : :|:

Qy   990 TACCGACGGCCGATGCTGCACCCCCCACAGAACCACCCACCCTGCCGGTGGAGTTCAAGTG 1049

Db   1024 nm 1025
          :

Qy   1050 CC 1051

RESULT 11
ID   Q10572 standard; DNA; 1047 BP.
AC   Q10572;
DT   09-APR-1991 (first entry)
DE   Human Natriuretic Peptide Receptor B.
KW   NPRB; ANP; BNP; CNP; kidney failure; heart failure; protein kinase;
KW   hyperaldosteronism; glaucoma; guanyl cyclase.
OS   Homo sapiens.
FH   Key   Location/Qualifiers
FT   Peptide   1..22
FT   /label= signal sequence
FT   Protein   12
FT   /label= mature NPBR
FT   Domain   23..455
FT   /label= extracellular domain
FT   /note= "binds natriuretic peptides A,B and C]"
FT   Domain   456..456
FT   /label= transmembrane domain
FT   Domain   479..1047
FT   /label= cytoplasmic domain
FT   /note= "GC and protien kinase activity"
FT   Modified -site 24..26
FT   /label= N-glycos_site
FT   Modified -site 35..37
FT   /label= N-glycos_site
FT   Modified -site 161..163
FT   /label= N-glycos_site
FT   Modified -site 195..197
FT   /label= N-glycos_site
FT   Modified -site 244..246
FT   /label= N-glycos_site
FT   Modified -site 277..279
FT   /label= N-glycos_site

FT   Modified -site 349..351
FT   /label= N-glycos_site
FT   Modified -site 600..602
FT   /label= N-glycos_site
PN   WO9100292-A.
PD   10-JAN-1991.
PF   22-JUN-1990; U03586.
PR   23-JUN-1989; US-370673.
PA   (GETH) GENENTECH INC.
PI   Chang M, Goeddel D, Lowe D;
DR   WPI; 91-036711/05.
DR   N-PSDB; Q10324.
PT   Natriuretic protein receptor B - for diagnosis and treatment of
PT   kidney failure, heart failure, hyperaldosteronism, glaucoma etc.
PS   Claim 3; Fig 1; 49pp; English.
CC   The sequence was derived from the DNA encoding natriuretic peptide
CC   receptor B, NPRB, having guanyl cyclase (GC) activity and protein
CC   kinase activity. The DNA can be inserted into expression vectors
CC   for the prodn. of the protein, opt. after being mutated to produce
CC   NPRB analogues. The protein has a mol wt. of 115 kD (calculated Mr=
CC   114,952). The protein (or variants) can be used in treatment of
CC   natriuretic peptide disorders, and also to isolate peptides using
CC   affinity chromatography. Antibodies with affinity for NPRB can
CC   also be prepd.
SQ   Sequence 1047 BP;   87 A;   15 C;   83 G;   51 T;

DB 2;   Score   79; Match 9.1%;   QryMatch 3.8%;   Pred. No. 9.36e-27;
Matches   86; Conservative 270; Mismatches 578; Indels 13; Gaps 12;

Db   85 vdnknyhdnhmnngngcvynaasvarnashwrnnmntagavasgnsakndhyrtnvrtgn 144
         ::: :: ::::   :: ::|:   : : | |: | :: :|| |::|:|||

Cp   970 CTCGGTATGTCTTCATGCTGGTGCAGCCAGAAGCTCAAACTTGATAGGCTTGGAGATTT 911

Db   145 sankngnnvvtnhghnnwtaraannynndartddrnhyntnngvnnangsnnsvnhnvya 204
          : | :: ::: :|| :| : ||:|::| | | :

Cp   910 TGGGAGTACGGATGCACTTTTCCCCTTCTTAATGTTCTCTTCCAGTCAGCTTCGCAAG 851

Db   205 rnngnnnathnnrangrnvynycngnnmnhmnnnanrnnntngdyrnynyndvngnsnra 264
          : | :: || | || |: ::| : | | | |:: :   : :|

Cp   850 GCCTGACCATGCACAGGCGCCTCTGCTTCTCTAGCCTGCAGGAGGCGCTTCATTGTAA 791

Db   265 gntatgrnwdnrtrnnananrnanntvnvtyrnnnvnynnnnrnnnrarndngvn 324
          : |: :||: | : : ::|: | | :: : :

Cp   790 CCCG--GGTGGAGATGCCCATCCCACAGGTCTTGGAACAGGCGCTCCACTCT-GTGGTCT 734

Db   325 ngnsmnnnagcnydgnnyanvnnntnnnggtrndgnrnvnkmngrryhgvtgnvvmdk 384
          | | : || || |:| || | :: | || :|::|| : :::

Cp   733 GGACCAGGCAG-TTGGCTCTAATCATAGTTGGTCTGGCCAAACGTGTCTTCCAGTCGG 675

Db   385 nndrntdnvnwamgdndsgdnnnaahysganknnwwtgrnnnwvkgannsdnnncandnd 444
          : : : :: || :: |:   : | | :: : || |:

Cp   674 TAGCCCCGAGGGCAGGCCCAACCACCGTTTGGTCCTTGGGCTCGTCACACACCCACTCC 615

Db   445 dnscdktnnstnanvangtgntnmmgvsnnnnrknmnnknnasmnwrnr-wnnnngn 503
          :: ||::   : | | : ::: |:

Cp   614 TCCAGGCATTTCCCGGCACCTTGACCCTCCTCGGAGAGGGCAGTCAGGGCTGGCCAGA 555

Db   504 snryhkgagsrntnsnrgsaygsmtahgkymnnantghnkgnvvankhvnkkrnnntrn 563
          : : :   : | ::: ||: :| |:|:: | | :: |:| ::: :|:|

Cp   554 CGAACGTCCATGCTGCACAGGGGCATGCAGCCCACCACCCCCGTTCCAGGCACGTGCACTG- 496

Db   564 vnnnkhmrdvmnhntrnngacndnnncnvtnycnrgsqndnnndannddwmnrysn 623
          : ::::: :| |: |: | :   ::| :|| |

Cp   495 GTACTTGCAGCTGCTCTGGA-AGGACTCTCCGCTGCGGTACACCGTACCACCGAAGATGC 437

```
Db  624 nndnvkgmannhnssnhshgsnkssncvvdsrnvkntdygnasnrstanndnnanyak 683
            : :::: ::  ::::: :::: ::::: :  :|::|  :: :::  |:
Cp  436 AGGGACCACCATCTTTGGCGG-TGCACGCCCATCTTGCGGTTGGCCGGGGAGCCGAAG 378

Db  684 knntannnsgnnnttgmaadvysgnnnmnanrsgnnynngndnsnknnvkvrng 743
            :  |  |||    ::::        :     :  ::
Cp  377 TCACAGAAGAGCCCTTGTCGGGTGCGCAGGGTCGCGTCGCAGCTCGCCCAGC 318

Db  744 nrnynrnsndrtnnnnvnnvnmmrcwandnanrndngnnkgnnrirnnkngtsnndnnn 803
            : : : :: :   : ::|   ::  :: : ::  :  :  : :
Cp  317 TGCTTGGCCAGACGCGGCAGCAGCCGCAGCCC-TCCAGCACCAGGCTCACGCCCCCGG 259

Db  804 rmnnyannmknvnnrtnaynnnkrkanannynnnhsvannnkrgntvnanandsvtny 863
            :: ::  :  :  |  ::     :|::     |:    : ::
Cp  258 GCAGCCGCGCGCCGGCTCGTCCGGGCACCCGACGCCCCGCTCCAGTTCTGCCCGACGGC 199

Db  864 nednvgntansan-stnmvvtnnndnytcndann-dnndvykvntngdaymv-vs-gnn 919
            :: :|  :|  :::  ::|   :| :|  |:|:::::  ::|::|:|  :|  :|
Cp  198 CGGCCCGGCTCCAGAGGCCAGGAGGACCACCAGAGCCACCGGACGGCCCCATACTGGC 139

Db  920 grngnhannarmananndavsnnrnrhndnnrrngvhtgnvcagvrgnkmnrycn 979
            |:|  :||          :  ::  ::  :::: | :| :|  ::  |
Cp  138 GGCGGTCATGGTTGGCACTCCGGCGGCCGGAGGCGGCCGGTGCCGCGAGCCGGAGCCG 79

Db  980 ngdtvntasrmsngnanknhvsttkdandngcnnnnnrrgdvnmk 1026
            |:  ::::     :|  ::|  :   |  :::::
Cp   78 CGGGGCCTGAGCGCTGCCGGTCGGAGCTGGGAGCCGGACGCGGCG 32
```

RESULT 12
ID Q57417 standard; cDNA to mRNA; 372 BP.
AC Q57417;
DT 19-OCT-1994 (first entry)
DE CYR61 like protein.
KW Human cDNA; library; enzyme; protein; ss.
OS Homo sapiens.
PN WO9403599-A.
PD 17-FEB-1994.
PF 04-AUG-1993; J01095.
PR 04-AUG-1992; JP-208077.
PR 13-NOV-1992; JP-327619.
PR 26-FEB-1993; JP-061431.
PA (SAGA) SAGAMI CHEM RES CENTRE.
PI Iwahori A, Kato S, Kato T, Kim N, Oh S, Sekine S;
DR WPI; 94-065688/08.
DR P-PSDB; R46078.
PT cDNA of human origin and proteins coded by it – which may be
PT expressed by in vivo or in vitro translation using sense RNA or
PT antisense DNA corresponding to the cDNA.
PS Claim 1; Page 32-33; 167pp; Japanese.
CC mRNA expressed in human fibrosarcoma cell line HT-1080 was
CC isolated and used to construct a cDNA library using vector
CC pKA1. Clone HP00021 encoding CYR61-like protein
CC was isolated.
SQ Sequence 372 BP;   93 A;   95 C; 107 G;   77 T;

DB 10; Score   67; Match 68.9%; QryMatch 3.2%; Pred. No. 2.18e-20; Indels 0; Gaps 0;
Matches 122; Conservative   0; Mismatches 55;

```
Db  70 taccaaaacgggaaagtttccagcccaactgtaaacatcagtgcacatgtattgatggc 129
           |||| | ||| |||  ||||||| ||||| || | ||||||||  |  || ||
Qy 457 TACCGGCAGCCGAGAGTCCTTCCAGCAGCTGCGCAGTTACCAGTCCACGTGCCTGGACGGG 516
```

```
Db 130 gccgtgggctgcattcctcgtgtccccaagaactatctcccaacttgggctgtccc 189
          || |||||||||| |||||| ||||| ||||   |  ||| |||| ||||| |||
Qy 517 GCGGTGGGCTGCTCCATGCCCCTGCGAGCATGGAGCGTGGCCTGCGCTCTGCCCGACTGCCCC 576

Db 190 aaccctcggctgcaaagttaccggcagtgcggtgcgcaggagtgggtctgtgacgag 246
          ||| || ||| | ||||| |||||| |||| | ||||||| |||||||||||||
Qy 577 TTCCCGAGGAGGTCAAGTCGCCCGGGAAATGCTGCGAGGAGTGGGTGTGTGACGAG 633
```

RESULT 13
ID Q36039 standard; cDNA; 216 BP.
AC Q36039;
DT 24-MAY-1993 (first entry)
DE Homologous to exon 3 of CTGF gene.
KW avian nephroblastoma; avian myeloblastoma virus;
KW stringent hybridisation; ss.
PN WO9300430-A.
PD 07-JAN-1993.
PF 25-JUN-1992; F00589.
PR 25-JUN-1991; FR-007807.
PA (CNRS) CENT NAT RECH SCI.
PI Martinerie C, Perbal B;
DR WPI; 93-036377/04.
DR P-PSDB; R31605.
PT Nucleotide sequences hybridising to regions of chicken nov gene –
PT useful as probes for detecting complementary sequences to
PT evaluate development and/or differentiation of tumours
PS Claim 14; Page 34; 67pp; French.
CC This sequence is homologous to exon 3 of the CTGF gene and was used
CC as a probe in Northern hybridisations with different human tissues
CC and cell lines. Probes homologous to exon 2 of the chicken nov gene
CC and to exon 4 of the CTGF gene (Q36035 and Q36040, respectively)
CC were also used in the hybridisation experiments. The results showed
CC that the human homologue of the chicken nov gene and the CTGF gene
CC belonging to the same family are expressed differently in different
CC tissues or lines as different RNA transcripts.
SQ Sequence 216 BP;   62 A;   56 C;   60 G;   38 T;

DB 6; Score   65; Match.67.2%; QryMatch 3.1%; Pred. No. 2.43e-19; Indels 0; Gaps 0;
Matches 127; Conservative   0; Mismatches 62;

```
Db   1 gcttacaggccagaagccaccctaggagtagaagtctgactcaagtgtcaactgcatt 60
          ||||| | ||||||| ||| | || |||  || || |||||||
Qy 670 GCTTACCGACTGGAAGACACGTTTGCCCAGACCCAACTATGATTAGAGCCAACTGCCTG 729

Db  61 gaacagaccagagtggacagcatgctccaagagctgtgtatggggttctccaccggg 120
          | |||||||||||| |||||| || ||||||||||| ||||| ||||||||||||
Qy 730 GTCCAGACCACCAGAGTGGACGCCGTGTTCCAAGACACTTGTGGGATGGCATCTCCACCCGG 789

Db 121 gtcaccaataggaaccgtcaatgtgagctgagatgctgaaacagactcggctcggtggtcgg 180
          || ||||||     |||      | | |  | |||||| || ||  |||||||||| ||
Qy 790 GTTACCCAATGACAACACGCCTCCTGCAGGCTAGAGAAGCAGAGCCGCCCTGCCATGGTCAGG 849

Db 181 ccctgtgaa 189
          || ||| |||
Qy 850 CCTTGCGAA 858
```

RESULT 14
ID Q36040 standard; cDNA; 435 BP.
AC Q36040;
DT 24-MAY-1993 (first entry)
DE Homologous to exon 4 of CTGF gene.
KW avian nephroblastoma; avian myeloblastoma virus;

```
*******************************************************
```

MPSRCH (TM)

```
*******************************************************
```

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelliGenetics, Inc.

MPsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Fri Sep 22 21:58:20 1995; MasPar time 1084.25 Seconds
                                           1117.032 Million cell updates/sec

Tabular output not generated.

Title: >US-08-167-628-1
Description: (1:2075) from 5408040.seq
Perfect Score: 2075
N.A. Sequence: 1 CCCGGCCGACAGCCCCGAGA.........AGGAAATGTGGTAGCTCACG 2075
    Comp: GGCCGGCTGTCGGGGCTCT.........TCCTTTACACCATCGAGTGC

Scoring table: TABLE default
                Gap 6

Nmatch STD :   Dbase 0; Query 0

Searched: 267821 seqs, 291840210 bases x 2

Database: embl-new6
   1 BCT
   2 EST1
   3 EST2
   4 EST3
   5 EST4
   6 EST5
   7 FUN
   8 INV1
   9 INV2
  10 MAM
  11 ORG
  12 PLN
  13 PRI
  14 PRO
  15 ROD
  16 STS
  17 SYN
  18 UNC
  19 VRT
  20 VIR

Database: genbank89
  21 BCT1
  22 BCT2
  23 BCT3
  24 BCT4
  25 BCT5
```

```
  26  BCT6
  27  INV1
  28  INV2
  29  INV3
  30  INV4
  31  INV5
  32  MAM1
  33  MAM2
  34  PAT1
  35  PAT2
  36  PHG
  37  PLN1
  38  PLN2
  39  PLN3
  40  PLN4
  41  PLN5
  42  PLN6
  43  PLN7
  44  PRI1
  45  PRI2
  46  PRI3
  47  PRI4
  48  PRI5
  49  PRI6
  50  PRI7
  51  PRI8
  52  PRI9
  53  ROD1
  54  ROD2
  55  ROD3
  56  ROD4
  57  ROD5
  58  ROD6
  59  ROD7
  60  STR
  61  STS1
  62  STS2
  63  STS3
  64  STS4
  65  SYN
  66  UNA
  67  VRL1
  68  VRL2
  69  VRL3
  70  VRL4
  71  VRL5
  72  VRL6
  73  VRT1
  74  VRT2
  75  VRT3

Database:   genbank-new6
  76  BCT
  77  EST1
  78  EST2
  79  EST3
  80  EST4
  81  EST5
  82  EST6
  83  EST7
  84  EST8
  85  INV
  86  MAM
  87  PHG
```

	41	39	1.9	3578 89	HSPLOD3	Human lysyl hydroxyla	9.36e-06
	42	39	1.9	609 64	HUMUT866	Human chromosome 2 ST	9.36e-06
c	43	39	1.9	328 63	HS196XB4	H. sapiens (D1S228) D	9.36e-06
	44	38	1.8	187 51	HUMRPTPOLE	Human repeat polymorp	3.16e-05
	45	38	1.8	192 48	HUMCHR8F	Human chromosome 8 DN	3.16e-05

ALIGNMENTS

```
RESULT     1
LOCUS        HUMCONGRO    2075 bp    mRNA            PRI        30-APR-1992
DEFINITION   Human connective tissue growth factor, complete cds.
ACCESSION    M92934 M36965
KEYWORDS     growth factor; mitogen.
SOURCE       Homo sapiens (library: lambda gt11) connective cDNA to mRNA.
ORGANISM     Homo sapiens
             Eukaryota; Animalia; Chordata; Vertebrata; Mammalia; Theria;
             Eutheria; Primates; Haplorhini; Catarrhini; Hominidae.
REFERENCE    1  (bases 1 to 2075)
AUTHORS      Bradham,D.M., Igarashi,A., Potter,R.L. and Grotendorst,G.R.
TITLE        Connective tissue growth factor: a cysteine-rich mitogen secreted
             by human vascular endothelial cells is related to the SRC induced
             immediate early gene product CEF-10
JOURNAL      J. Cell Biol. 114, 1285-1294 (1991)
MEDLINE      91373462
COMMENT      NCBI gi: 180923
FEATURES             Location/Qualifiers
     source          1..2075
                     /organism="Homo sapiens"
                     /cell type="endothelial"
                     /sequenced_mol="cDNA to mRNA"
                     /tissue_type="connective"
                     /tissue_lib="lambda gt11"
     mRNA            1..2075
     sig_peptide     130..192
                     /codon_start=1
     CDS             130..1179
                     /note="NCBI gi: 180924"
                     /codon_start=1
                     /product="growth factor"
                     /translation="MTAASMGPVRVAFVVLLALCSRPAVGQNCSGPCRCPDEPAPRCP
                     AGVSLVLDGCGCCRVCAKQLGELCTERDPCDPHKGLCDFGSPANRKIGVCTAKDGAP
                     CIFGGTVYRSGESFQSSCKYQCTCLDGAVGCMPLCSMDVRLPSPDCPFPRRVKLPGKC
                     CEEWVCDEPKDQTVVGPALAAYRLEDTFGPDTMIRANCLVQTTEWSACSKTCGMGIS
                     TRVTNDNASCRLEKQSRLCMVRPCEADLEENIKKGKKCIRTPKISKPIKFELSGCTSM
                     KTYRAKFCGVTDGRCCTPHRTTTLPVEEKCPDGEVMKKNMMFIKTCACHYNCPGDND
                     IFESLYRKMYGDMA"
     mat_peptide     193..1176
                     /codon_start=1
                     /product="growth factor"
BASE COUNT      491 a    558 c    546 g    480 t
ORIGIN

DB 48;  Score 2075;  Match 100.0%;  QryMatch 100.0%;  Pred. No. 0.00e+00;
Matches 2075;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Db    1  cccggccgacagcccgagacgacgagcccggcgcgtcccacctccgaccaccg   60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    1  CCCGGCCCAGCAGCCCGAGAGACGAGCCCGGCGCGTCCCGGTCCCACCTCCGACCACCG   60

Db   61  ccagcgctccaggccgcccgctcccgctcgccaccgcccctccgctccgcccgca  120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   61  CCAGCGCTCCAGGCCGCCCGCTCCCCGCTCGCCCCACCGCCCCTCCGCTCCGCCCGCA  120
```

```
88  PLN
89  PRI
90  ROD
91  STS
92  STR
93  SYN
94  UNA
95  VRL
96  VRT
```

Database: u-embl43_89
            97 ALL

Statistics: Mean 12.417; Variance 7.139; scale 1.739

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Match	Query Length	DB	ID	Description	Pred. No.
	1	2075	100.0	2075	48	HUMCONGRO	Human connective tiss	0.00e+00
	2	2075	100.0	2075	52	S56201	connective tissue gro	0.00e+00
	3	2022	97.4	2312	44	HSCTGF	H.sapiens mRNA for co	0.00e+00
	4	921	44.4	2330	57	MUSTGFB	Mouse mRNA sequence.	0.00e+00
	5	909	43.8	2267	55	MUSFISP12B	Mouse FISP-12 protein	4.05e-224
	6	354	17.1	4128	55	MUSFISP12A	Mouse FISP-12 protein	9.46e-159
	7	264	12.7	1976	74	GGNOVMRNA	G.gallus nov mRNA.	4.22e-150
	8	252	12.1	1927	73	CJU13063	Coturnix japonica cel	8.80e-62
	9	127	6.1	1805	73	CHKCEF	Chicken CEF-10 protei	1.39e-35
	10	88	3.3	2018	55	MUSCYR61A	Mouse Cyr61 mRNA, com	9.25e-23
	11	68	3.3	1469	46	HSNOVH5	H.sapiens novH gene e	6.81e-21
	12	65	3.1	435	46	HSNOVH4	H.sapiens novH gene e	4.80e-19
	13	62	3.0	452	46	HSNOVH3	H.sapiens novH gene e	7.98e-18
	14	60	2.9	5196	53	MMCYR61G	Mouse growth factor i	4.80e-13
	15	52	2.5	557	46	HSNOVH12	H.sapiens novH gene e	1.79e-08
c	16	44	2.1	226	86	BTMSAT90	B.taurus microsatelli	1.79e-08
	17	44	2.1	226	10	BTMSAT90	B.taurus microsatelli	6.41e-08
c	18	43	2.1	261	63	HS154XC7	H. sapiens (D1S216) D	2.27e-07
c	19	42	2.0	221	33	SHPMAF4P	Ovis aries dinucleoti	7.93e-07
c	20	41	2.0	322	63	HUM4STS367	Human chromosome 4 se	7.93e-07
	21	41	2.0	392	33	SSMIC	S.scrofa pseudogene c	7.93e-07
c	22	41	2.0	499	63	HUM4STS439	Human chromosome 4 se	2.74e-06
	23	40	1.9	184	53	MMD3NDS3	M.musculus D3nds3 mic	2.74e-06
c	24	40	1.9	365	63	HS345WH9	H. sapiens (D4S1617)	2.74e-06
c	25	40	1.9	340	62	HS116XG1	H. sapiens (DXS986) D	2.74e-06
c	26	40	1.9	163	62	GGU23939	Gorilla gorilla dinuc	2.74e-06
c	27	40	1.9	260	79	R72445	yj90h07.s1 Homo sapie	2.74e-06
c	28	40	1.9	353	63	HS248WC5	H. sapiens (D2S172) D	2.74e-06
	29	40	1.9	146	51	HUMVTR14K	Human variable tandem	2.74e-06
c	30	40	1.9	1782	55	MUSATP4B1	Mouse gastric (H+,K+)	2.74e-06
c	31	40	1.9	418	33	SS13N06R	S.scrofa genomic DNA	2.74e-06
c	32	40	1.9	397	32	BTMSINRBN	B.taurus microsatelli	2.74e-06
c	33	39	1.9	669	47	HSU19800	Human m2 muscarinic a	9.36e-06
c	34	39	1.9	3578	13	HSLOD3	Human lysyl hydroxyla	9.36e-06
c	35	39	1.9	387	63	HUM4STS237	Human chromosome 4 se	9.36e-06
c	36	39	1.9	354	51	HUMSIRPOBN	Human simple repeat p	9.36e-06
	37	39	1.9	3578	46	HSPLOD3	Human lysyl hydroxyla	9.36e-06
	38	39	1.9	529	64	HUMUT5046	Human STS UT5046.	9.36e-06
	39	39	1.9	407	64	HUMUT454	Human chromosome 8 ST	9.36e-06
c	40	39	1.9	8633	30	PFPFMDR1	P.falciparum pfmdr1 g	9.36e-06

```
Db  1081  atgttcatcaagacctgtgctgctgctgccattacaactgtcccggacaatgacatctttgaa  1140
          ||||||||||||||||| ||| |||||||||||||||||||||||||||| |||||||||||||
Qy  1081  ATGTTCATCAAGACCTGCCTGCCATTACAACTGTCCCGGACACAATGACATCTTTGAA       1140

Db  1141  tcgctgactacaggagatgtacgagacatgcatgaggcagcagagtgagagacatt        1200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1141  TCGCTGACTACACGGAGACGATGTACGAGAGATGCCATGAGGCAGCAGAGTGAGACACATT    1200

Db  1201  aactcattagactggaacttgaactgattcacatctcatttttccgtaaaatgatttca    1260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1201  AACTCATTAGACTGGAACTTGAACTGATTCACATCTCATTTTTCCGTAAAAATGATTTCA    1260

Db  1261  gtagcacaagttatttaaatctgttttctcaactggggaaaagattccaaccaattca    1320
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1261  GTAGCACAAGTTATTTAAATCTGTTTTCTCAACTGGGGAAAAGATTCCCACCCAATTCA     1320

Db  1321  aaacattgtgccatgtcaaacaaatagtctatcttcccagacaactggtttgaagaatgt   1380
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1321  AAACATTCTGCCATGTCAAACAAATAGTCTATCTTCCCCGACACTGGTTTGAAGAATGT     1380

Db  1381  taagacttgacagtggagggtacggccggccggtagtatcatcagatcgactcttatacgag   1440
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1381  TAAGACTTGACAGTGGAGGGTACCGGGTTAGTATGATCATCAGATCGACTCTTATACGAG     1440

Db  1441  tttaggagcagtggggtgtaattgagaaggaaaatttagcgtgctcactgacct        1500
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1441  TTTAGGAGCAGTGGGAGTGTAATTGAGAAGGAAAATTTTAGCGTGCTCACTGACCT        1500

Db  1501  taatatgctgctctatttgaagtgaagttgtaattgagaaggaaaatttagcgtgctcactgacct  1560
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1501  TAATATGCTGCTATTTGAAGTGTAATTGAGAAGTGAATTGATTTTAGCGTGCTCCTGACCT    1560

Db  1561  gcctgctagcccagtgacagctaggatggtgcattctccagccatcaagagactgagtcaa    1620
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1561  GCCTGTAGCCCCAGTGACGACTAGGATGTGCATTCTCCAGCCATCAAGAGACTGAGTCAA     1620

Db  1621  gttgttccttaagtcagaacacagacagctcagctgacattctgattcgaatgacactgt   1680
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1621  GTTGTTCCTTAAGTCAGAACAGCAGACTCAGCTCTGACATTCTGATTCGAATGACACTGT    1680

Db  1681  tcaggaatcggaatcctgtcgattagactggacagcttgtggcaagtgaattgcctgta    1740
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1681  TCAGGAATCGGAATCCTGTCGATTAGACTGGACAGCTGTGGCAAGTGAATTGCCTGTA     1740

Db  1741  acaagccagatttttaaaattatattgtaatattgtgtgtgtgtgtgtgtata          1800
          ||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Qy  1741  ACAAGCCCAGATTTTTAAAATTTTATATTGTAAATTTATATTCTGAACACCATAGGTAGAAT   1800

Db  1801  tatatatatatgtacagttatctaagttaatttaagttgtcttttttatttt          1860
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1801  TATATATATATGTACACATTTAAGTTAATTAATTTAAGTTGTTTTGTTGTTTGTTTATTTT     1860

Db  1861  tgtttttaatgctttgatatttcttgctgctcaattctgaacaccataggtagaat        1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1861  TGTTTTTAATGCTTTGATATTTCAATGTTCAATTCTGAACAACCATAGGTAGAAT         1920

Db  1921  gtaaagcttgctgatcgtcaaagcatgaacatgataactatatggaaattctgctcag       1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1921  GTAAAGCTTGTCGTCGTCTCATCGTTCAAAGCATGAAATGGATACTTATATGAAATTCTGCTCAG  1980

Db  1981  ataagaatgcatcctgaaaacagattgtttgcaaagggaggaggcatcagtgtcttggca    2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1981  ATAGAATGACAGTCCGTCAAAACAGATTGTTTGCAAAAGGGGAGGGGGAGGCCATCAGTGTCTTGGCA  2040
```

```
Db  121   gtgccaaccatgaccaccgccgccagtatggaccccgtccgatcgcctcggtgtcctcctc  180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  121   GTGCCAACCATGACCACCGCCGCCGCCAGTATGGGCCCCCAGTGTCCGTGTCCGTCCTCTC    180

Db  181   gccctcgagcggccgcggcggcgtcggccagaactgcgagcgagcggagcgtgccggac     240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  181   GCCCTCTGCCAGCCGCAGGCGCCGCCCGGCGTCGCGGGCCCGTGCCGGCCGGTGCCCGGAC    240

Db  241   gagccggcgcgcggcggcgcgcgcggggcgtgagcctcgtgacgacgtcgcggctgctgc    300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  241   GAGCCGGCGCCGCGGCGCGCGCGCGGGGCGTGCCCTCGTGACGCCGGCGCGGCTGCTGC     300

Db  301   cgcgtctgcgccaagcagctgggcagctgtgtgcagcgagcggcgacccctgcgacccgcac  360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  301   CGCGTCTGCGCCAAGCAGCTGGGCAGCTGTGTGCAGCGGCGACCCCGACCCCGACCCCCAC   360

Db  361   aagggcctcttctgtgactcggctcccggcccccgccccgaagatccaagatcggcggtgcaccgcc  420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  361   AAGGGCCCTCTTCTGTCTCTCCGGTGGTACGGCGGGTGTCATCGGCGTCGTGTCGCACCGCC    420

Db  421   aaagaatggtgctccctgacgtcacgtaccggcgccggtggctgggtcggtgcatgcctgtgc   480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  421   AAAGAATGGTGCTCCCTGACGTCTTCGCATCTTCGGTGGTACCGGTGTCATGCCCTCTGC     480

Db  481   agcagctgcaagtaccagtgcacgtgcacgtgccgctgaccgtggcggggtcggtgggtcggtgccc  540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  481   AGCAGCTGCAAGTACCAGTGCACGTGCACGTGCCGGCGTGGCTGGGTGGGGTCGGTGGGTCGGTGCCC  540

Db  541   agcatgacgttcgtctgccagccctgactgcctcccgaggaggtcaagctgccc        600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  541   AGCATGACGTTCCTGCCAGCCCTGACTGCCTCCCCGAGGAGGTCAAGCTGCCC          600

Db  601   gggaaatgctgcgaggagtgggtgtgtgacgagccagccagccaaaacccctggttgggcct    660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  601   GGGAAATGCTGCGAGGAGTGGGTGTGTGACGCCGGCCAAGGACCCAAACCGTGGTTGGGCCT   660

Db  661   gccctcgcggcttaccgactggaagaacacgttggcccagaccaactatgattagagcc     720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  661   GCCCTCGCGGCTTACCGACTGGAAGAACACGTTCGCCCAGACCACTATGATTAGAGCC       720

Db  721   aactgcctggtccagacacacagagtggagcgcgcctgttccaagacctgtgggatgggcatc  780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  721   AACTGCCTGGTCCAGACACCACAGACTGGAGCGCCCTGTTCCAAGACCTGTGGGATGGCATC   780

Db  781   tccaccgggtaccaatgacaacgcctctgcaggctagaagcagagccgcctgtgc      840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  781   TCCACCGGGTACCAATGACAACGCCCTCCTGCGGCTAGAAGCAGAGAGCCGCCTGC       840

Db  841   atggtcaggcctgcaagctcacagctcaagtttgagcttctggctgcacccagatgaag     900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  841   ATGGTCAGGCCTGCAAGCTCACAGCTCAAGTTTGAGCTTCTGGCTGCACCAGCATCAAG     900

Db  901   cgtactcccaaaatctccaagccatcaagtttgagcttctggctgcacccagatgaag      960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  901   CGTACTCCCAAAATCTCCAAGTTTGAGCTTTCTGGCTTTCGGCTTTCAATTCACCATCAAG    960

Db  961   acataccgactaaattctgtgaagtatgtacccgacggcgccatgctgcaccccccacaga   1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  961   ACATACCGACTAAATTCTGTGGAGTATGTACCGAGTATGTACCCGACGCCGCCACCCCCACAGA  1020

Db  1021  accacaccctgccggtcaagtcaagtcaaagtcaaagtcccatgaagagagaacatg       1080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1021  ACCACACCCTGCCGGTGACTCAAGTCAAGTCGCATGAAGAGAGAACAACATG           1080
```

```
Db  2041 ggctgatttctaggtaggaaatgtggtagctcacg 2075
          |||||||||||||||||||||||||||||||||||
Qy  2041 GGCTGATTTCTAGGTAGGAAATGTGGTAGCTCACG 2075

RESULT      2
LOCUS       S56201       2075 bp    mRNA              PRI       29-APR-1993
DEFINITION  connective tissue growth factor [human, umbilical vein endothelial
            cells, mRNA, 2075 nt].
ACCESSION   S56201
KEYWORDS    .
SOURCE      human umbilical vein endothelial cells.
  ORGANISM  Homo sapiens
            Unclassified.
REFERENCE   1  (bases 1 to 2075)
  AUTHORS   Igarashi,A., Bradham,D.M., Okochi,H. and Grotendorst,G.R.
  TITLE     Connective tissue growth factor
  JOURNAL   Journal of Dermatology 19 (11), 642-643 (1992)
  MEDLINE   93187114
REMARK      GenBank staff at the National Library of Medicine created this
            entry [NCBI gibbsq 127024] from the original journal article.
            This sequence comes from Fig. 1.
COMMENT     NCBI gi: 266234
FEATURES             Location/Qualifiers
     source          1..2075
                     /organism="Homo sapiens"
                     /note="human"
     CDS             130..1179
                     /note="Description: connective tissue growth factor, CTGF;
                     Method: conceptual translation supplied by author. This
                     sequence comes from Fig. 1.  CTGF; NCBI gi: 266235"
                     /codon_start=1
                     /product="connective tissue growth factor"
                     /translation="MTAASMGPVRVAFVVLLALCSRPAVGQNCSGPCRCPDEPAPRCP
                     AGVSLVLDGCGCCRVCAKQLGELCTERDPCDPHKGLFCDFGSPANRKIGVCTAKDGAP
                     CIFGGTVYRSGESFQSSCKYQCTCLDGAVGCMPLCSMDVRLPSPDCPFPRRVKLPGKC
                     CEEWVCDEPKDQTVVGPALAAYRLEDTFGPDPTMIRANCLVQTTEWSACSKTCGMGIS
                     TRVTNDNASCRLEKQSRLCMVRPCEADLEENIKKGKKCIRTPKISKPIKFELSGCTSM
                     KTYRAKFCGVCTDGRCCTPHRTTTLPVEEKCPDGEVMKKNMMFIKTCACHYNCPGDND
                     IFESLYRKMYGDMA"
BASE COUNT      491 a    558 c    546 g    480 t
ORIGIN

DB 52;   Score  2075;  Match 100.0%;  QryMatch 100.0%;  Pred. No. 0.00e+00;
Matches  2075;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Db    1 ccccggccgcagacccgagacgcagcccggcccggcccgtcccacctccgaccaccg 60
         ||||||||| || ||||||||||||||||||| ||||||||||| | | ||||||||||
Qy    1 CCCCGGCCACAGCCCCGAGACGCAGCCCGGCCGGCCCGTCCCCGGCTCCCGACCACCG 60

Db   61 ccagcgctccaggccgccccggctcccccgctcgccgcaacgcgcctccgtccgccgca 120
         |||||||||||||||||||||| ||||||||||||||| ||||||||||||||| |||
Qy   61 CCACCGCTCCAGGCCGCCCCGGCTCCCCCGCTCGCCGCCACCGCGCCTCCGTCCCCCGCA 120

Db  121 gtgccaaccatgaccggccgccgcccagctggcccgctatggccccgctcggtcctcctc 180
         |||||||||||||||||||||||||| | ||| |||| |||||||| | ||||||||||
Qy  121 GTGCCAACCATGACCGGCCGCCGCCCGGCCCAGTATGGGCCCGTCGTTCGGTCCTCCTC 180

Db  181 gccctctgcagccggccgcctcggccagaaactgcagcggccgccgtgccggtccggac 240
         |||||||||||||||||| || |||||||| ||||||||||| || ||||||| |||||
Qy  181 GCCCTCTGCAGCCGGCCGGCCGTCGGCCAGAACTGCAGCGGCCCGTGCCGGTGCCCGGAC 240

Db  241 gagccggcgccgcgctgcccggccggggtgagcctcgtgctgggacggctgcggctgtgc 300
         ||||||||||
```

```
Db  241 GAGCCGGCGCCGCGCTGCCCGGCCGGGGTGAGCCTCGTGCTGGACGGCTGCGGCTGCTGC 300
         ||||||||||
Qy  241 GAGCCGGCGCCGCGCTGCCCGGCCGGGGTGAGCCTCGTGCTGGACGGCTGCGGCTGCTGC

Db  301 cgggtcgcggccaagcagctgggcgagctgtgcaccgagcgagacccctgcgaccccac 360
         ||| |||||||||||| || |||||||||||||| || ||||| |||||||||| |||||
Qy  301 CGGGTCTGCGCCAAGCAGCTGGGCGAGCTGTGCACCGAGCGAGACCCCTGCGACCCCCAC 360

Db  361 aagggcctcttctgtgacttcggctcccgccaaccgcaaagatcggcgtgcaccgcc 420
         |||||||||| || ||||| ||||| ||||| |||||||| ||| ||||||||||||
Qy  361 AAGGGCCCTTCTGTGACTTCGGCTCCCGCCAACCCCAAGATCGGCGTGTGCACCGCC 420

Db  421 aaagatgtgctccctgcatcttcggtggtacggtaccgcagcagagagtcctccag 480
         |||||||||||||||||||||||||||||| |||| | | |||| | |||||||||
Qy  421 AAAGATGTGCCCTGCACTGCTTCGGTACGGTACCGCAGCAGAGAGTCCTTCCAG 480

Db  481 agcagctgcaagtaccagtgcacgtgcctgacggggcgtgggctcgatgccccgtgc 540
         |||||||||||||||||||||||||||||||||||||| ||||| |||||||| ||||
Qy  481 AGCAGCTGCAAGTACCAGTGCACGTGCCTGGACGGGGCGTGGGCTGGCTGCATGCCCCTGTGC 540

Db  541 agcagtgacgtgtctgccagcctgactgcccttcccgaggaggtcaagctgcc 600
         ||||| ||||||| ||||||||| |||||||| | |||||||||||||||||||
Qy  541 AGCATGGACGTTGTCTCGCCAGCCTGACTGACCCTTCCCGAGGAGGTCAAGCTGCCC 600

Db  601 gggaaatgctgcgaggagtggtgtgacgagccaaggaccaaaccggtggttgggcct 660
         ||||||||||||||||||||||||||||||||||||||||||||| || |||||||| |
Qy  601 GGGAAATGCTGCGAGGAGTGGGTGTGACGAGCCAAGGACCAAACCGGTGGTTGGGCCT 660

Db  661 gccctcggcggttaccgactggaagacacgttggccagaccacctactatgattagagc 720
         ||||| ||||||||||||||||||||||||||| ||||||||||||||| || |||||||
Qy  661 GCCCTCGCGGCTTACCGACTGGAAGACACGTTTGGCCGACACCCAACTATGATTAGAGCC 720

Db  721 aactgctggtccagacacagagtggagcgcctgttccagacctgtgggatggcatc 780
         |||||||||| ||||| ||||| ||||| ||||||||||| ||||||||||||||||
Qy  721 AACTGCCTGGTCCAGACAGACCACCAGAGCTGAGCGCGCTGGTTCCAGACCTGTGGGATGGCATC 780

Db  781 tcaacccgggttaccaatgacaacgcctctccaggctagagaagcagagcgcctgtgc 840
         ||||||| || ||||||||||||||||||||||| |||||||||||||||||||||||||
Qy  781 TCCACCCGGGTTACCAATGACAACGCCTCTCCAGGCTAGAGAAGCCAGAGCGCCCTGTGC 840

Db  841 atggtcaggcctgcgaagctgacctggaagagaacattaagagggcaaaaagtgcatc 900
         ||||| |||||||| ||||| |||||||| ||||||| ||||| |||||| |||||||
Qy  841 ATGGTCAGGCCTTGCGAAGCTGACCTGGAAGCAGAACATTAAGAAGGCAAAAACTGCATC 900

Db  901 cgtactcccaaaatctccaagcctatcaagttgagctttcggctgccaccagcatgaag 960
         ||| ||| |||||||||||| || |||| ||| ||| |||| || ||||| ||||||||
Qy  901 CGTACTCCCAAAATCTCCAAGCCTATCAAGTTTGAGCTTTCTGGCTGCCACCAGCATGAAG 960

Db  961 acataccgagctaaattctgtggagttgtaccagacgcccatgctgcacccccccacaga 1020
         |||||||||||| ||| ||| |||||| |||||| | || ||| | |||||||||||||||
Qy  961 ACATACCGAGCTAAATTCTGTGGAGTATGTACCGACGGCCCATGCTGCACCCCCCACAGA 1020

Db 1021 accacacctgccggggagttcaagtgccgtcgagtcatgagagagaacatg 1080
         |||||||||| || |||||| |||| | | ||| ||||| |||| || |||
Qy 1021 ACCACCACCCTGCCGGTGGAGTTCAAGTGTCGACGGCCGGCAGTCATGAAGAAGAACATG 1080

Db 1081 atgttcatcaagacctgctctgccattacaactgtccggagacaatgacatcttgaa 1140
         |||||||||||| || ||| |||||||| || || |||||| |||||||||| ||||||
Qy 1081 ATGTTCATCAAGACCTGCGCTGCCATTACAACTGCCCGGGAGACAATGACATCTTTGAA 1140

Db 1141 tcgctgctactacaggaagatgtacggagacatgcatggagcatgacgtgagagacatt 1200
         ||||| ||||| |||||| |||||||| ||||| ||||||| | ||||| ||||| || |
Qy 1141 TCGGTGCTACTACAGGAGATGTACGGAGACATGCATGGAGCATGAGAGTGAGAGACATT 1200

Db 1201 aactcattagactggaacttgaactgattcacatcctcattttccgtaaaatgatttca 1260
         ||||||||||||
```

```
Qy  1201  AACTCATTAGACTGCGAACTGAACTGATTCACATCTCATTTTCCGTAAAAATGAATTCA 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1261  gtagcacagttatttaaatctgtttctcaactggggaaagattcccaccaattca 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1261  GTAGCACAGTTATTTAAATCTGTTTTCTCAACTGGGGAAAAGATTCCCACCCAATTCA 1320

Db  1321  aaaattgtgcatgtcaaacaaatagtctatcttcccagacactggtttgaagaatgt 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1321  AAACATTGTGCCATGTCAAACAAATAGTCTATCTTCCCCAGACACTGGTTTGAAGAATGT 1380

Db  1381  taagacttgacagtggaactacattagtacacagcaccagaatgtatattaaggtgtggc 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1381  TAAGACTTGACAGTGGAACTACATTAGTACACAGCACCAGAATGTATATTAAGGTGTGGC 1440

Db  1441  tttaggagcagtgggagggtaccggccggttagtatcatcagatcctttcttatacgag 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1441  TTTAGGAGCAGTGGGAGGGTACCGGCCGGTTAGTATCATCAGATCCATTCTTATACGAG 1500

Db  1501  taatatgcctgctatttgaagtgtaattgagaaggaaaattttagcgtgctcactgacct 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1501  TAATATGCCTGCTATTTGAACTGTAATTGAGAAGGAAAATTTTAGCCTGCTGCCACTGACCT 1560

Db  1561  gcctgtagcccagtgacgagctaggatgtgcattctccagccatcaagagactgagtcaa 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1561  GCCTGTAGCCCCACTGACCAGCTAGGATGTGCATTCTCCAGCCATCAAGACAGACTGAGTCAA 1620

Db  1621  gttgttccttaagtcagaacagcagacctcagctctgaatcgaatgacactgt 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1621  GTTGTTCCTTAAGTCAGAACAGCAGACACTCAGCTCTGATTCGAATGACACTGT 1680

Db  1681  tcaggaatcggaatcctgtcgattagactggacagcttgtggcaagtgaattgcctgta 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1681  TCAGGAATCGGAATCCTGTCGTTGGTCGATTAGACTGGACAGCTTGTGGCAAGTGAATTTGCCTGTA 1740

Db  1741  acaagccagattttttaaaatttatattgtaaatattgtgtgtgtgtgtata 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1741  ACAAGCCAGATTTTTTAAAATTTTTTATATTGTAAATATTGTGTGTGTGTGTGTATA 1800

Db  1801  tatatatatatgtacagttatctaagttaaagtgttttgtgcctttttatttt 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1801  TATATATATATCTACAGTTATCTCAAGTTAAAGTTTTAAAGTTGTTTGTCGCCTTTTATTTT 1860

Db  1861  tgttttaatgctttgatatttcaatgttagcctccaattctgaacaccatagtagaat 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1861  TGTTTTAATGCTTTGATATTTCAATGTTAGCCTCCAATTTCTGAACACCATAGCTAGCAAT 1920

Db  1921  gtaaagctgctgtcgatcgttcaaagcatgaaagtgatacttatatggaaattctgctcag 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1921  GTAAAGCTTGTCTGCTGTCATGTCTGCATGCAGTGCCATCTATATGGAAATTCTGCTCAG 1980

Db  1981  atagaatgacagtccgtcaaaacagattgtttgcaaaggggaggcatcagtgcttggca 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1981  ATAGAATGACACGTCCCGTCAAAAACAGATTGTTTGCAAAGGGGAGGCATCAGTGTCTTGGCA 2040

Db  2041  ggctgatttctagggtaggaaatgtggtagctcacg 2075
          ||||||||||||||||||||||||||||||||||||
Qy  2041  GGCTGATTTCTCAGGTAGGTAGGAAAATGTGGTAGCTCACG 2075
```

RESULT   3
LOCUS   HSCTGF   2312 bp   RNA   PRI   25-APR-1994
DEFINITION H.sapiens mRNA for connective tissue growth factor.
ACCESSION X78947
KEYWORDS   connective tissue; growth factor.

```
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa/Eumycota group;
            Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata;
            Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Choanata;
            Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates;
            Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 2312)
AUTHORS     Oemar,B.S., Werner,A., Yang,Z., Garnier,J.M., Gentz,R. and
            Luescher,T.F.
TITLE       Differential cloning and expression of human connective tissue
            growth factor
JOURNAL     Unpublished
REFERENCE   2 (bases 1 to 2312)
AUTHORS     Oemar,B.S.
TITLE       Direct Submission
JOURNAL     Submitted (22-APR-1994) to the EMBL/GenBank/DDBJ databases. B.S.
            Oemar, University Hospital Basel, Dept of Research, Lab of Vascular
            Research, Hebelstr. 20, 4031 Basel, SWITZERLAND
COMMENT     NCBI gi: 474933
FEATURES             Location/Qualifiers
     source          1..2312
                     /organism="Homo sapiens"
                     /isolate="patient C.F."
                     /tissue_type="aortic, (blood vessel)"
                     /chromosome="6"
                     /map="6q23.1"
                     146..1195
                     /note="NCBI gi: 474934"
                     /codon start=1
     CDS             /product="connective tissue growth factor"
                     /translation="MTAASMGPVRAFVVLLALCSRPAVGQNCSGPCRCPDEPAPRCP
                     AGVSLVLDGCGCCRVCAKQLGELCTERDPCDPHKGLFCDFGSPANRKIGVCTAKDGAP
                     CIFGGTVYRSGESFQSSCKYQCTCLDGAVGCMPLCSMDVRLPSPDCPFPRRVKLPGKC
                     CEEWVCDEPKDQTVVGPALAAYRLEDTFGPDTMIRANCLVQTTEWSACSKTCGMGIS
                     TRVTNDNASCRLEKQSRLCMVRPCEADLEENIKKGKKCIRTPKISKPIKFELSGCTSM
                     KTYRAKFCGVCTDGRCCTPHRTTLPVEEKCPDGEVMKKNMMFIKTCACHYNCPGDND
                     IFESLYRKMYGDMA"
     mat_peptide     224..1192
                     /product="connective tissue growth factor"
     misc_feature    1480
                     /note="a-g transition"
     misc_feature    1486..1487
                     /note="a-c transversion"
     misc_feature    1505
                     /note="a-c transversion"
     polyA_signal    2262..2267
BASE COUNT    572 a    594 c    585 g    561 t
ORIGIN

DB 44;  Score 2022;  Match 99.5%;  QryMatch 97.4%;  Pred. No. 0.00e+00;
Matches 2068;  Conservative 0;  Mismatches 4;  Indels 7;  Gaps 5;

Db    17  ccgggccgacagcccgagacgacagccgcggcgccgtcccacctcgaccaccg 76
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     1  CCCGGCCGACAGCCCCGAGACGACAGCCCGCGGCGCCGTCCCACCTCCACCACCG 60

Db    77  ccagcgctcaggccgccggccccgtccgccgccaccgcgcctccggtcgcccgca 136
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    61  CCAGCGCTCCAGGCGCCGGCCCGCTCCCCGGCTGCCGCCACCGCGCCCTCCGGCCCGCA 120

Db   137  gtgcaaccatgaccggccgccgcatatggccccgtccgctgccttcgtcctcctc 196
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   121  GTGCCAACCATGACCGCCGGCCAGTATGGGCCCCGTCGGCTGCCTTCGTCCTCCTC 180
```

```
Db  197  gccctctcgcagcagcggccggccgtcggccagaactgagccggccgttgcccggcgac  256
         |||||||||||||||||||||||||||||||||||||||||||
Qy  181  GCCCCTCTGCAGCCGCGGCCGGCCGTCGGCCAGAACTGAGCCGGCCGTTGCCCGGCGAC  240

Db  257  gagccgcgccgcggcccgcccggcgggcgtctgcctctgctgctgcggctgcgctgctgc  316
         |||||||||||||||||||||||||||||||||||||||||||||
Qy  241  GAGCCGGCGCCGCGGCCCGCCCGGCGGGCGTCTGCCTCTGCTGCTGCGGCTGCGCTGCTGC  300

Db  317  cggtctcggccaagcagctgggcgagctgcaccagcggcaccctgcgaccgcac  376
         ||||||||||||||||||||||||||||||
Qy  301  CGGTCTCGGCCAAGCAGCTGGGCGAGCTGCACCAGCGGCACCCTGCGACCGCAC  360

Db  377  aagggcctcttctgtgacttcggctcccggccaaccgaacatcggcctgtgcaccgcc  436
         |||||||||||||||||||||||||||||||||
Qy  361  AAGGGCCTCTTCTGTGACTTCGGCTCCCGGCCAACCGAACATCGGCCTGTGCACCGCC  420

Db  437  aaagatggtgctccctgcatcttcggtggtacggtgccgcagcggagagtccttccag  496
         |||||||||||||||||||||||||||||||
Qy  421  AAAGATGGTGCTCCCTGCATCTTCGGTGGTACGGTGCCGCAGCGGAGAGTCCTTCCAG  480

Db  497  agcagctgcaagtaccagtgacgtgcctgctgacaggcggtgggctgcatgccccgttgc  556
         ||||||||||||||||||||||||||||||
Qy  481  AGCAGCTGCAAGTACCAGTGACGTGCCTGCTGACAGGCGGTGGGCTGCATGCCCCGTTGC  540

Db  557  agcatgacagttcgtctgccaggccctgactgccctccgagagggtcaagctgccc  616
         |||||||||||||||||||||||||
Qy  541  AGCATGACAGTTCGTCTGCCAGCCACTGCCCTCCGAGGAGGGTGGGCTGCAAGCTGCC  600

Db  617  gggaaatgctgcgaggagtgggtgtgacgagccaagaacaacctggttgggcct  676
         |||||||||||||||||||||||||||||
Qy  601  GGGAAATGCTGCGAGGAGTGGGTGTGACGACGCCCAAGGACCAAACCTGGTTGGGCCT  660

Db  677  gccctgcggcttaccgactggaagacacgtttggccagccaactatgattagagcc  736
         ||||||||||||||||||||||||||||
Qy  661  GCCCTGCGGCTTACCGACTGGAAGACACGTTTGGCCAGCCAACTATGATTAGAGCC  720

Db  737  aactgctggtccagacagagtggagcggcctgttccaagacctgtggatgggcatc  796
         |||||||||||||||||||||||||||||
Qy  721  AACTGCTGGTCCAGACAGAGTGGAGCGGCCTGTTCCAAGACCTGTGGGATGGCATC  780

Db  797  tccacccgggttaccaatgacaacgcctctcgagctagaagagagcccgcctgtgc  856
         ||||||||||||||||||||||||||||||
Qy  781  TCCACCCGGGTTACCAATGACAACGCCCTCTCGAGGCTAGAAGAGCACCGCCTGTGC  840

Db  857  atggtcaggcctttgacctgacctgaagagaacattaagaaggggcaaaaagtgcatc  916
         ||||||||||||||||||||||||||||
Qy  841  ATGGTCAGGCCTTTGACCTGACCTGGAAGAAGCAGACATTAAGAAGGGCAAAAAGTCCATC  900

Db  917  cgtactccgaagatggttgtctggttctgctcgcaccagcatgaag  976
         ||||||||||||||||||||||||||||
Qy  901  CGTACTCCCAAAATCTCCAAGCCTATCAAGTTTGAGCTTTCTGCTGCCACCAGCATGAAG  960

Db  977  acataccgagctaaattctgtggagtatgtaccgacggccgatgtgcaccccccacaga  1036
         |||||||||||||||||||||||||||||||
Qy  961  ACATACCGAGCTAAATTCTGTGGAGTATGTGAGCATGTACCGACGGCCGATGTGCACCCCCCACAGA  1020

Db  1037 accacaccctgccggtggagttcaagtccgttatggcctgagtcatgaagaagaacatg  1096
         |||||||||||||||||||||||||||||||||
Qy  1021 ACCACACCCTGCCGGTGGAGTTCAAGTCCGTCAGTGCCAGGTCATGAAGAAGAACATG  1080

Db  1097 atgttcatcaagacctgtgctgctccaactgtcccggacaatgacatctttgaa  1156
         ||||||||||||||||||||||||||||||
Qy  1081 ATGTTCATCAAGACCTGTGCCTGCCATTACAACTGTCCCGGACAATGACCATCTTTGAA  1140
```

```
Db  1157 tcgctgtactacaggaagatgtacggagacatggcatgaagcagaagagtgagaagacatt  1216
         |||||||||||||||||||||||||||||||||||||
Qy  1141 TCGCTGTACTACAGGAAGATGTACGGAGACATGGCATGAAGCAGAAGAGTGAGAAGACATT  1200

Db  1217 aactcattagactggaactgaactgattcacatctccattttccgtaaaaatgatttca  1276
         ||||||||||||||||||||||||||||||||||
Qy  1201 AACTCATTAGACTGGAACTGAACTGATTCACATCTCCATTTTTCCGTAAAAATGATTTCA  1260

Db  1277 gtacacaagttattaaatctgtttttctaactggggaaaagattcccaccccaattca  1336
         ||||||||||||||||||||||||||||||||||
Qy  1261 GTAGCACAAGTTATTAAATCTGTTTTTCTAACTGGGGAAAAGATTCCCACCCAATTCA  1320

Db  1337 aaacattgtgccatgtcaaacaaatagtctatcttcccagacactggtttgaagaatgt  1396
         ||||||||||||||||||||||||||||||||
Qy  1321 AAACATTGTGCCATGTCAAACAAATAGTCTATCTTCCCAGACACTGGTTTGAAGAATGT  1380

Db  1397 taagacttgacagtggaactacattagtacacagcaccagaatgtatattaaggtgtggc  1456
         |||||||||||||||||||||||||||||||||||||||
Qy  1381 TAAGACTTGACAGTGGAACTACATTAGTACACAGCACCAGAATGTATATTAAGGTGTGGC  1440

Db  1457 tttaggagcagtggagggtaccagcagaaagttagtatcatcagatag-ctcttatac  1515
         |||||||||||||||||||||                      ||||||||||
Qy  1441 TTTAGGAGCAGTGGAGGGTACCGGAGGGTACCGGGCCC---GGTTAGTATCATCAGATCGACTCTTATAC  1497

Db  1516 gagtaatatgctgctatttgaagtgtaattgagaagaaatttagcgtgctcactga  1575
         ||||||||||||||||||||||||||||||||
Qy  1498 GAGTAATATGCCCTGCTATTTGAAGTGTAATTGAGACTGGAGGAAAATTTAGCCTGCTGCTCACTGA  1557

Db  1576 cctgcctgtagcccagtgacgacgctaggatgtgcattctccagcatcaagagactgagt  1635
         ||||||||||||||||||||||||||||||||||||
Qy  1558 CCTGCCTGTAGCCCCAGTGACGATGCCCAGTGCCATTCTCCAGCATCAAGAGACTGAGT  1617

Db  1636 caagtgtccttaagtcagaacagcagactcagctctgattcgaatgcac  1695
         ||||||||||||||||||||||||||||||
Qy  1618 CAAGTTGTTCCTTAAGTCAGAACAGCAGACTCAGCTCTGACATTCTGATTCGAATGACAC  1677

Db  1696 tgttcaggaatcggaatcctgtcgattagactgacagctgtgtggcaagtgaattt-cct  1754
         ||||||||||||||||||||||||||||||
Qy  1678 TGTTCAGGAATCGGAATCCTGTCGATTAGACTGGACTGGCAAGTTGTGGCAAGTTGGGCCTT  1737

Db  1755 gtaacaagcagattttttaaaatttatatgtaatattgtgtgtggctgtgt  1814
         ||||||||||||||||||||||||||||
Qy  1738 GTAACAAGCCAGATTTTTTAAAATTTATATTTTAAAATTATATGTAAATATTGTGTGTGGCTTTTTAT  1797

Db  1815 atatatatatatgtacagtgttatctaagttaattaaagaaggtaattgtgctttttat  1874
         |||||||||||||||||||||
Qy  1798 ATATATATATTATATATGTACAGTGTATCATCGTAAGTTAATTAAAGTGTTTGTGCCTTTTTAT  1857

Db  1875 ttttgtttttaatgcttgatattcaatgttgacctcaattctgaacaccatagtag  1934
         |||||||||||||||||||||||||||
Qy  1858 TTTTGTTTTTAATGCTTGATATTCAATGTGACCTCAATTCTCGAACACCATAGGTAG  1917

Db  1935 aatgtaaagctgtctgatcgttcaaaacatgaaatgatacttatatgaaattct-ct  1993
         |||||||||||||||||||||||||||
Qy  1918 AATGTAAAGCTTGTCTGATCGTTCAAAGCATGAAATGATACTTATATGGAAATTCTGCT  1977

Db  1994 cagatagaatgacatcgccgtcaaaaacagattgtttgcaaggggagcatcagtgtcctt  2053
         |||||||||||||||||||||||||
Qy  1978 CAGATAGAATGACAATGCCGTCAAAAACAGATTGTTTGCAAAGGGGAGGCATCAGTGTCACG  2036

Db  2054 ggcaggctgattctaggtaggaaatgtgtagctcacg  2092
         |||||||||||||||
Qy  2037 GGCAGGCTGATTTCTAGGTAGGAAATGTGTAGCTCACG  2075
```

```
                                                                                    469
Qy   410  TGTCCACCCCAAAGATGGTGCTCCCTGCATCTTCGGTTGGTTACGGTTACGGCTACCGCAGCCGGAG  469

Db   541  agtccttccaaagcagctgcaaataccaatgcacttgcctgatggggccctggggctgcg       600
          |||||||||   |||||||||  ||| ||||||| |||||||||||||||||||||||||||
Qy   470  AGTCCTTCCAGAGCAGCTGCAAGTACCAGTCGCACGTGCCTGGACGGGGCCGGTGGCTGCA       529

Db   601  tgcccctatgcagcatgacgtgcgcctgccagccgacctgactgcgccctcccgagaaggg     660
          ||||||| ||||||||| ||||  ||| |||||||||||||||||||||||||||| ||||
Qy   530  TGCCCCTGTGCAGCATGACGTTCGTCTGCCCCAGCCCGACTGCGCCCTTCCCGAGGAGGG     589

Db   661  tcaagctgctgggaaatgctgcgagagtggtggtgtgacgagccaaggaccgcacag        720
          |||||||||   ||||||||||||||||||||||||||||||||||||||| ||
Qy   590  TCAAGCTGCCGGGAAATGCTGCGGAGAGTGGTGGTGCTGTCGACGAGCCCAAGGACCAAACCC     649

Db   721  cagttggccctgcctagctgctaccgactggagacaacatttgcccagaccaacta        780
          |||| ||||||| || ||||||||||||| |||||||||||||||||||||||||||
Qy   650  TGGTTGGGCCTGCCCTCGCCGGTTACCGACTGGAGCACCGTTGCCCAGACCCAACTA     709

Db   781  tgatgcggcaactgctggtccagaccacagtggagcggctgttctaagacctgtg        840
          ||||  |||||||||||||||||||||||||||||||||||||| ||||||||||
Qy   710  TGATTAGGCCAACTGCCTCTGTCCAGACCACCAGTGGAGCGGCCTGTTCCAAGACCCGTG     769

Db   841  gaatggcatctccaccgagttaccaatgacaatacctctgcagactggagaagcaga     900
          | |||||||||||||||| |||||||||| ||||||||||||  ||||||| |||||
Qy   770  GGATGGGCATCTCCACCGGGGTTACCAGCCGGGCCCTCTGCAGCGTAGCAGAAGCAGA     829

Db   901  gccgcctcgcatggtcaggccctgcgaagctgacctggaggaaacattaagaagggca     960
          ||||||||||||||| |||||||||||||||||||||||||||| |||||||||||||
Qy   830  GCCGCCTGTGCGATGGTCAGCGTTGCAGGCCTTGCCAAGCTGACCTGGAAGAAGGAACATTAAGAGGGCA     889

Db   961  aaaagtgcatccggaccacctaaatcgccaagcctgcaagtttgagcttctggctgca    1020
          |||||| ||||||||||||||   |||||||||||||||||||||||||||||||||
Qy   890  AAAAGTGCATCCGTACTCCCAAAATCTCCAAGCCTATCAAGTTTGAGTTTCTGGCTGCA     949

Db  1021  ccagtgtgaagacatacagggctaagttctgcggggtgtgcacagacgccgctgctgca    1080
          |||| ||||||| |||||| |||| ||| || || ||||||||||||||||||||||
Qy   950  CCACCATGAAGAACATGATGTTCATCAAGAAGAACCTGTGCCTGCCATTACAACTGTCCCCGGAGCAATG    1009

Db  1081  caccgcacagaaccaccactctgccagtggagttcaaatgcccgatggcggagatcatga    1140
          || |||||||||||||||||||| |||||||||||||| ||||||||||| ||||||||
Qy  1010  CACCGCACAGAACCACCACCCTGCCCGTGAGTTCAAGTTCAAGTGCCCTGACGGCGAGGTCATGA    1069

Db  1141  aaaagaatatgatgttcatcaagacctgctcctacaactgtctggggacaatg        1200
          ||||| |||||||||||||||| ||||| ||| ||  ||||||||||||||
Qy  1070  AGAAGAACATGATGTTCATCAAGACCGTGCCTGCCATTACAACTGTCCCCGGAGCAATG    1129

Db  1201  acatctttggtccctgctacgaggacgtcagcggctaaagccaggaag        1260
          |||||||||||||||| |||||||||||||||||||||| |
Qy  1130  ACATCTTTGAATCGTCGTACTACAGGAAGATGATGGCATGGACAGGCCAGAGAG    1189

Db  1261  taaggacacgaactcattagactataacttgaactgagttgcatctcatttctctgt    1320
          ||||||||| ||||||||||||||  |||||||||||||||||||   |||||
Qy  1190  TGAGAGACATTAACTCATTAGACTGGAACTTGAACTGTTGAACTGGCAT        1247

Db  1321  aaaaacaattacagtagcacat-taatttaaatctgtgttttaactaccgtggaggaa    1379
          |||||| |||||||| | ||| ||||||||||| ||||||||||||
Qy  1248  AAAAATGATTTCAGTAGCACAAGTTATTTAAATCTGTTTTCTAACTG--G-GGGAAAAG    1304

Db  1380  ctatcccacaagtgagaacgttatgtcatggccataccaagtgtcaacctcaga      1439
          | ||||||||||||||||||||| |||||||||||||||||||||||||||||
Qy  1305  AT-TCCCACCCAATTCAAACCATTGTGCCATGTCAA-ACAAATAGTCTATCTTCCCAGA    1362

Db  1440  cactggttcgagacagtttacacttgacagtgt--t-cattagcgcacagtgccgaa    1496
          |||||||| |  |||||||||||||  |  |||||||||    |||||
Qy                                                                (1440)
```

```
RESULT       4
LOCUS        MUSTGFB      2330 bp     mRNA        ROD      01-NOV-1991
DEFINITION   Mouse mRNA sequence.
ACCESSION    M80263
KEYWORDS     transforming growth factor-beta.
SOURCE       Mus musculus cDNA to mRNA.
ORGANISM     Mus musculus
             Eukaryota; Animalia; Chordata; Vertebrata; Mammalia; Theria;
             Eutheria; Rodentia; Myomorpha; Muridae; Murinae.
REFERENCE    1  (bases 1 to 2330)
AUTHORS      Brunner,A., Chinn,J., Neubauer,M.G. and Purchio,A.F.
TITLE        Identification of a gene family regulated by transforming growth
             factor-beta
JOURNAL      DNA Cell Biol. 10, 293-300 (1991)
MEDLINE      91229699
COMMENT      NCBI gi: 201945
FEATURES     Location/Qualifiers
   source    1..2330
             /organism="Mus musculus"
             /cell_line="AKR-2B"
             /sequenced_mol="cDNA to mRNA"
   mRNA      1..2330
   CDS       204..1250
             /note="NCBI gi: 201945"
             /codon_start=1
             /translation="MLASVAGPISLALVLLALCTRPATGQDCSAQCQCAAEAAPHCPA
             GVSLVLDGGCGCCRVCAKQLGELCTERDPCDPHKGLFCDFGSSPANRKIGVCTAKDGAPC
             VFGSVYRSGESFQSSCKYQCTCLDGAVGCVPLCSMDVRLPSPDCPFPRRVKLPGKCC
             EEWVCDEPKDRTAVGPALAAYRLEDTFGPDPTMMRANCLVQTTEWSACSKTCGMGIST
             RVTDNTFCRLEKQSRLCMVRPCEADLEENIKKGKKCIRTPKIAKPVKFELSGCTSVK
             TYRAKFGVCTDGRCCTPHRTTTLPVEFKCPDGEIMKKNMMFIKTCACHYNCPGDNDI
             FESLYRKMYGDMA"
   sig_peptide  204..278
             /codon_start=1
   mat_peptide  279..1247
             /codon_start=1
   polyA_signal 2291..2296
   polyA_site    2318
BASE COUNT     589 a     598 c     595 g     548 t
ORIGIN
```

```
DB 57;    Score    921;   Match 80.8%;    QryMatch 44.4%;   Pred. No. 0.00e+00;
Matches 1544;  Conservative    0;  Mismatches 317;  Indels  51;  Gaps  29;

Db   241  tgtgtgctcctgcgccctctgcacccggctacgggcaggaactgcagccgcaatgtc     300
          ||||  ||||||||||||||||  ||| |||||| |||||||||| || |||  |
Qy   170  TGGTCCTCCTCCGCCCTCTGCAGCGCGGCCGGCGGCCCGGCGGCCTCGGCCGGCCGGCCGGCGCCGGCCGTGCC     229

Db   301  agtgcgcagcgaagcagcgccgcactgccgcccgcccgcgccgctggacggct        360
          ||||| | ||| | |||||| |||  | ||||||||||||||||| |||||||||
Qy   230  GGTGCCCCGGACGAGCCGCCGCCGCCCGCCGCCCGCTGGCCGGCCTCGTGCTGGCACGGCT     289

Db   361  gcgggctgctgccgcgctgcgccaagcagctgggagaactgtacggagcgtgaccct     420
          ||||||||||||||| |||||||||||||||||| || |||||||  ||||||| ||
Qy   290  GCGGGCTGCTGCCGCGTCTGCGCCAAGCAGCTGGGCGAGCTGTGCACCGAGCGCGACCCT     349

Db   421  gcgaccacacaaggcgccttctgcgatttcggctcccccgcgccaaccgcaagattggag     480
          ||||| |||||||||||| ||||||||| || || |||||||  |||  ||||| | |
Qy   350  GCGACCCCGCACAAGGGCGCCTTCTGCGACTTCGGCTCCTCGCCCGCCAACCGCAAGATCGGCG     409

Db   481  tgtgcactgcaagatggtgcacccgctgtcttcggtggtcggttcggtgtaccgcagcgtg     540
          ||||||  |||||||||||||||||  ||||| |||||||||||||||||||||||||||  |
Qy                                                                      (481)
```

```
source          1..2267
                /organism="Mus musculus"
                /sub_species="domesticus"
                /sequenced_mol="cDNA to mRNA"
                /tissue_type="liver/kidney"
CDS             138..1184
                /gene="fisp-12"
                /map="A3-B1"
                /note="NCBI gi: 193316"
                /codon_start=1
                /product="FISP-12 protein"
                /translation="MLASVAGPISLALVLLALCTRPATGQDCSAQCQAEEAPHCPA
                GVSLVLDGCGGCRVCAKQLGELCTERDPCDPHKGLFCDFGSPANRKIGVCTAKDGAPC
                VFGGSVYRSGESFQSSCKYQCTCLDGAVGCVPLCSMDVRLPSPDCPFPRRVKLPGKCC
                KEWVCDEPKDRTAVGPALAAYRLEDTFGPPTMMRANCLVQTTEWSACSKTCCGMGIST
                RVTNDNTFCRLEKQSRLCMVRPCEADLEENIKKGKKCIRTPKIAKPVKPFELSGCTSVK
                TYRAKFCGVCTDGRCCTPHRTTLPVEFKCPDGEIMKKNMMFIKTCACHYNCPGDNDI
                FESLYRKMYGDMA"
polyA_signal    2224..2230
                /gene="fisp-12"
                /map="A3-B1"

BASE COUNT      570 a    569 c    582 g    546 t
ORIGIN          chromosome 10.

DB 55;  Score   909;  Match 81.9%;  QryMatch 43.8%;  Pred. No. 0.00e+00;
Matches 1353;  Conservative   0;  Mismatches 270;  Indels 29;  Gaps 19;

Db   175 tgtgctcctcgccctgcaacccgcctgctacggccaggactgcagcgcgcaatgtc 234
         ||||||||||| ||||||| |||||||| |||||||||| ||||||   |||||
Qy   170 TGGTCCTCCTCGGCCCTGCAGCCCGGCCGGCCGGCGGCCGGGCCGCCGTGCC 229

Db   235 agtcgcagccgagcagcgccgcactgccccgccccgccggcgtgagcgtcggtgctggacggct 294
Qy   230 GGTCCCGGACGACCCGGCGTGCCCGGCGGCGTGAGCCTCGTCGCTGACGGCT 289

Db   295 gcggctgctgccgcgtgcggccaagcagctgggagaactgtgtacggagcgtgacccct 354
Qy   290 GCGGCTGCTGCCGCGGTGCGGCCAAGCAGCTGGGCGAGCTGTGCACCGAGCGCGACCCCT 349

Db   355 gcgaccacacaaggcgcctctctgcgatttcggctccccgccgcaacccgcaagattggag 414
Qy   350 GCGACCCGCACAAGGGCCTCTTCTGCGACTTCGGCTCCCCGGCCAACCGCAAGATCGGCG 409

Db   415 tgtgcactgccaaagatggtgcacctgtgtcttcgttcggtgggtcggtgtaccgcagcggtg 474
Qy   410 TGTCCACCGCCAAAGATGGTGCCTCCCGCATCTTCGTTCGGTGTACGGTGTACCGCAGCGGAG 469

Db   475 agtccttccaaagcagctgcaaataccaatgcacttgcctggatgggccgtgggctgcg 534
Qy   470 AGTCCTTCCAGAGCAGCTGCAAGTACCAGTGCACGTGCCTGGACGGGGCCGGTGGGCTGCA 529

Db   535 tgccccatgcagcagcatggagcgcgctgccagcccgaccgacctgcccttcccgagaaggg 594
Qy   530 TGCCCCTGTGCAGCATGGACGTTGTCTGCGCCAGCCCGACTCCCCGTTCCCAGGAGGGG 589

Db   595 tcaagctgcctgggaaatgctgcaaggagtgggtgtgtgacgaccgcacag 654
Qy   590 TCAAGCTGCCCGGGAAATGCTGCAAGGAGTGGGTGTGTGACGAGCCCAAGGACAGGACCAAACCG 649

Db   655 cagttggcctggcctgcctaccgactggaagacacatttggcccagaccaacta 714
Qy   650 TGGTTGGGCGTGCCCGCCCTGCGGCTTACCGACTGGAAGAGACACGTTTGGCCCAGAGACCCCACTA 709

Db   715 tgatgcgagccaactgccctggtccagaccaacgagtggagcgcctgttctaagaccctgtg 774
```

```
Qy  1363 CACTGCGTTTGAAGAATGTTAAGACTTGACAGTGACTACTACAGTACAGCACCAGA 1422

Db  1497 cgcacactgaggtgagtcctcctggaacagtggagtgccagagaagagaacaggtac 1556
         ||| || ||  || |||  |||||||   |||||||  |||  ||| |||
Qy  1423 TGTATATTAAGGTGTGGCTTAGGACGAGTGG--GAGGGTACCGGCCCG---GTTAG-TAT 1477

Db  1557 tagctgaggttatttaaaagcagcagtggcctactttttggagtgtaacaggaggg 1616
         | ||| |||| |  | |||||   |||  |    ||||||| |||||
Qy  1478 CATCAGATCGACTCTTATACG-AGTAATATGCCTGCTATTGAAGTGTRATTGAGAAGGA 1536

Db  1617 aaattaatagcatgcttgcagacagacctg---ctctagcggagagctgagca-tgtgtcct 1672
         |||| ||   |  ||| ||   |   ||||   |||| ||| |||  || |
Qy  1537 AAATTTAGGGTGCTACTCACTGACCTGCCTGAGCCCCAGTGACAGCTAGGATGTGCATTCT 1596

Db  1673 cca-ctag-atgaggctgagtccagctgttcttaag----aacagcagttcagctctg 1726
         ||| |||  |  |  ||||||| |  |||||| ||||    ||||| |||| |||||||
Qy  1597 CCAGCCCATCAAGAGACTGAGTCAAGTTGTTCCTTAAGTCAGAACACCAGACTCAGCCTCTG 1656

Db  1727 accattctgattccagtgacacttgtcaggagtcagagccctgtgtcttgttagactggacag 1786
         || |||| ||||| ||| ||||| |||||| ||   |  ||| | ||| ||||||||||
Qy  1657 AC-ATTCGATTCGAATGACACTGTTCAGGAATCGGAATCCTCGATTAGACTGGACAG 1715

Db  1787 cttgtggcaagtaagttgcctgtaacaagccagattttt---attgatattgtaaata 1842
         ||||||||||| | ||| |||| ||||||| |||||||   ||| ||||||||||||
Qy  1716 CTTGTGGCAAGTGAATTTGCCCTGTAACAAGCCCAGATTTTTTAAAATTTATATTGTAAATA 1775

Db  1843 tgtg-----ga-tatatatatatatatatttgtacagttatctaagttaattt 1896
         ||||     || |||||||||||||| || |  ||| ||||| |||| |||||
Qy  1776 TTGTCGTGTGTCGTGTGTGTATATATATATATATGTCAAGTTATCCAAGTTAATTT 1835

Db  1897 aaagtcatttgtt--tttgt-tttaagtgcttt-tgg--gatttttaaactgatagcctc 1949
         ||||| ||||||   ||| | ||| |||| ||| | |  ||||| ||| ||| |||||||
Qy  1836 AAAGTTGTTTGTGCCCTTTTATTTTTGTTTTAATGCTTTGATATTTCAAATGTTAGCCCTC 1895

Db  1950 aaactccaaaacaccataggtaggaacagaagcttatctg-tgatcaaaaca---aaggag 2006
         || ||| ||||||||||||||||||| | |||||| |||| ||  |   |||   |||
Qy  1896 AATTTCTGAACACCATAGGTAGGAAATGTAAAGCTTGTCTGATGCTTGTCTCAAAGCATGAAATGG 1955

Db  2007 atactgcagtgggaattgtgacctgagtgagtcagaagactctctgtcagaacaaacaaatgctgtgca 2066
         ||||| | || |||||| | |  |||| |||  ||||||   || ||||||||||| ||||| | ||
Qy  1956 ATACTTATATGGAAATTCTGCTCAGATAGAATGCAGCAGTCCGTCAAAAACAGATTGTTTGCA 2015

Db  2067 --ggtgataaagctatgtattggaagtcagacagatttctag-taggaaaatgtggt 2115
           ||| || || | | |  ||||||||  |   |||||||| ||||||||||||||
Qy  2016 AAGGGGAGGAGGCATCAGTGTCTTGCCAGGCTGATTTCTCAGGTAGGAAAATGTGGT 2067
```

```
RESULT  5
LOCUS       MUSFISP12B   2267 bp    mRNA            ROD       09-JUL-1991
DEFINITION  Mouse FISP-12 protein (fisp-12) mRNA, complete cds.
ACCESSION   M70642
KEYWORDS    cysteine-rich protein; growth factor-inducible gene.
SOURCE      Mus musculus (sub_species domesticus) liver/kidney cDNA to mRNA.
ORGANISM    Mus musculus
            Eukaryota; Animalia; Chordata; Vertebrata; Mammalia; Theria;
            Eutheria; Rodentia; Myomorpha; Muridae; Murinae.
REFERENCE   1  (bases 1 to 2267)
AUTHORS     Ryseck,R.-P., MacDonald-Bravo,H., Mattei,M.-G. and Bravo,R.
TITLE       Structure, mapping and expression of fisp-12, a growth factor
            inducible gene encoding a secreted cysteine-rich protein
JOURNAL     Cell Growth Differ. 2, 225-233 (1991)
MEDLINE     91363290
COMMENT     NCBI gi: 193315
FEATURES    Location/Qualifiers
```

```
           ||||  |||||||||||||||||||||||  |||||||  | |||||||  ||||||||||
Qy     710 TGATTAGAGCCAACTGCCTGGTCCAGACCACGAGTGGACGGCCGCCTGTTCCAAGACCTGTG 769

Db     775 gaatgggcatccacccggagttaccaatgacaatacctctgcagactggagaagcaga 834
           |||||||||||||| |||| ||||||||||||  |||| || |||||||||||||||
Qy     770 GGATGGCCATCTCCACCCGGGTTACCAATGACACACCCTCCTGCAGGCTAGAGAAGCAGA 829

Db     835 gccgcctctgcatggtcaggaccctgcgaggccgtgacctgagagaaaacattaagaaggca 894
           |||||||| ||||||| |||||| || |||| |||  |||||||||||||||||||||||  |
Qy     830 GCCGCCTGTGCATGGTGCAGGCCTTGCGAAGCTGACCTGGAGAGAACATTAAGAAGGCGA 889

Db     895 aaagtgcatccggacacctaaaatgcacctgtcaagtttgagctttctggctgca 954
           ||||||||||| | || ||| ||||| ||||||||||||||||||||||||||||
Qy     890 AAAAGTGCATCCGTACTCCCAAAATCTCCAAGCCTATCAGTTTGAGCTTTCTGGCTGCA 949

Db     955 ccagtgtgaagacatacagggctaagttctgcgggggtgtgcacagacgcgcgctgctgca 1014
           |||| |||||||||| | || ||| |||||||| |||||| ||||||  || ||||||||||
Qy     950 CCAGCGATGAAGACATACCGAGCTAAATTCTGTGGAGTATGTACCGACGGCGATGCTGCA 1009

Db    1015 caccgcacagaaccaccactctgccagtggagttcaaatcgcccatggccgatcatga 1074
           | |||||||||||||||| ||| |||||||||||||||||||  ||| |||||||||||
Qy    1010 CCCCCACAGAACCACCACCCTCCCGTGCGAGTTCAAGTCCCCTGACGGCGAGGTCATGA 1069

Db    1075 aaaagaatgatgttcatcaagacctgtgcctgccattacaactgtcctgggacaatg 1134
           | |||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Qy    1070 AGAGAACATGATGTTCATCAAGACCTGTGCCTGCCATTACAACTGTCCCGAGACAATG 1129

Db    1135 acatcttggtccctgtactacaggaaagatgtacggagacatggcgtaaagccaggaag 1194
           ||||||| |  |||||||||||||||| |||||| |  ||| |||||  |||||||| ||
Qy    1130 ACATCTTTGAATCGCTGTACTACAGGAAGAGATGTACCGGACACATGGCGATCAGCCAGCAGAG 1189

Db    1195 taaggggacgacgactcattagactataacttgaactgagttgcatctcattttcttctgt 1254
           |  || ||| ||| |||||||||| |||||||||||| || |||||  ||||||||  ||||
Qy    1190 TGAGACAGCATTAACTCATTAGACTGAACTGAACTGATTCACATCTCATTTT--TCCGT 1247

Db    1255 aaaaacaattacagtagcacat-taatttaaatctgtgttttaactaccgtgggaggaa 1313
           ||||| |||| |||||||||||  |||||||||||| |||||| |||| ||  | |||||
Qy    1248 AAAAATGATTTCAGTAGCACAAGTTATTTAAATCTGTTTTTCTAACTG--G-GGGAAAAG 1304

Db    1314 ctatccaccaaagtgagaacgttatgtcatggccataacaagtagtctgtcaacctcaga 1373
           || ||| ||| |||||| |||| |||||||||| ||| ||| ||||| || | ||||||||
Qy    1305 AT-TCCCACCCATTCAAAACATTGTGCCCATGTCAA-ACAAATAGTCTATCTTCCCCCACA 1362

Db    1374 cactggtttcggagacagttacacttgacagttgt---t-cattagcgcacagtgccagaa 1430
           |||||| ||| |||| ||| |||||||||||||    |||||||  |||||| |||||||
Qy    1363 CACTGTGTTTGAAGAATGTTAAGACTTGACAGTGGAACTACATTAGTTACACAGCACCAGA 1422

Db    1431 cgcacactgaggtgagtccctggaacagtggagatgccaggagagaaagaaagacaggtac 1490
           ||  | ||||| || || |||||| ||||| ||| || |||||||||||||| |||||
Qy    1423 TGTATATTAAGGTGTGGCTTTAGGACGCAGTGG-GAGGGTACCGGCCCCG---GTTAG-TAT 1477

Db    1491 tagctgaggttattttaaagcagcagtgctgcctactttttggagtgaaccggggaggg 1550
           ||  |||| | |  ||||||||||||||  |||| || | |||| ||||  |||||| |
Qy    1478 CATCGACATCCACTCTTATACG-AGTAATATGCCTGCTATTTGAAGTGTAATTGAGAAGGA 1536

Db    1551 aaattatagcatgctgcagacagacctg---ctctcagcgagagctgagca-tgtgtcct 1606
           ||||||||||||| |||| || ||| ||  |||| ||| || |||||||  || |||
Qy    1537 AATTTTAGCGTGCTCACTGACCTGCCTGTAGCCCCCAGTGACAGCTAGGATGGCATTCT 1596

Db    1607 cca-ctag-atgagggtgagtccagctgttcttaaga-acagcagtttcagccctctgac 1663
           |||  ||| ||||| | |||||| ||||||| ||||| |||||||| ||||| ||||| |
Qy    1597 CCAGCCATCCAAGAGACTGAGCTGAGTCAAGTTGTTCCTTAAGTCAGAACACAGCGACTCAGCTCTG 1656

Db    1664 -cattctgattccagtgacacttgtcaggagtcaggagccttgctcgtgtagactggacagc 1722
```

```
           |||||||||||  |||||| | |||||||||||||  ||| |  |||| |||| ||||||||||
Qy    1657 ACAITCTGAITCGAATGACACTGTTCAGGAATCGGAATCCTGTCGATTAGACTGGACAGC 1716

Db    1723 ttgtggcaagtaagtttgcctgtaacaagccagatttt----attgatattgtaaatat 1778
           |||||||||||| |||| |||||||||||||||||||||||    ||| |||| |||||| |
Qy    1717 TTGTGGCAAGTGAATTGCCTGTAACAAGCCAGATTTTTAAAAATTTATTATTGTAAATAT 1776

Db    1779 tgtggtatatatatatattattttgta 1810
           |||| |||||||||||||| |||| |||
Qy    1777 TGTGTGTGTGTGTGTGTATATATATATA 1808
```

```
RESULT          6
LOCUS           MUSFISP12A    4128 bp    DNA              ROD        09-JUL-1991
DEFINITION      Mouse FISP-12 protein (fisp-12) gene, complete cds.
ACCESSION       M70641
KEYWORDS        cysteine-rich protein; growth factor-inducible gene.
SOURCE          Mus musculus (sub species domesticus) liver/kidney DNA.
ORGANISM        Mus musculus
                Eukaryota; Animalia; Chordata; Vertebrata; Mammalia; Theria;
                Eutheria; Rodentia; Myomorpha; Muridae; Murinae.
REFERENCE       1  (bases 1 to 4128)
AUTHORS         Ryseck,R.-P., MacDonald-Bravo,H., Mattei,M.-G. and Bravo,R.
TITLE           Structure, mapping and expression of fisp-12, a growth factor
                inducible gene encoding a secreted cysteine-rich protein
JOURNAL         Cell Growth Differ. 2, 225-233 (1991)
MEDLINE         91363290
COMMENT         NCBI gi: 193313
FEATURES        Location/Qualifiers
     source          1..4128
                     /organism="Mus musculus"
                     /sub_species="domesticus"
                     /sequenced_mol="DNA"
                     /tissue_type="liver/kidney"
     TATA_signal     798..804
     exon            833..1114
                     /gene="fisp-12"
                     /map="A3-B1"
                     /number=1
     mRNA            join(833..1114,1202..1424,1675..1926,2059..2270,
                     2642..4006)
                     /gene="fisp-12"
                     /map="A3-B1"
     CDS             join(1052..1114,1202..1424,1675..1926,2059..2270,
                     2642..2938)
                     /gene="fisp-12"
                     /map="A3-B1"
                     /note="NCBI gi: 193314"
                     /codon_start=1
                     /product="FISP-12 protein"
                     /translation="MIASVAGPISLALVLLALCTRPATCQDCSAQCCQAEAAPHCPA
                     GVSLVLDGCCGCCRVCAKQLGELCTERDPDPHKGLFCDFGSPANRKIGVCTAKDGAPC
                     VFGGSVYRSGESFQSSCKYQCTLDGAVGCVPLCSMDVRLPSPDCPFPRRVKLPGKCC
                     KEWVCDEPKDRTAVGPALAAYRLEDTFCPDPTMMRANCLVQTTEWSACSKTCGMGIST
                     RVTNDNTFCRLEKQSRLCMVRPCEADLEENIKKGKKCIRTPKIAKPVKFELSGCTSVK
                     TYRAKFCGVCTDGRCCTPHRTTLPVEFKCPDGEIMKKNMMFIKTCACHYNCGDNDI
                     FESLIYRKMYGDMA"
     intron          1115..1201
                     /gene="fisp-12"
                     /map="A3-B1"
                     /number=1
     exon            1202..1424
```

Page 19:

```
                /gene="flsp-12"
                /map="A3-B1"
                /number=2
     intron     1425..1674
                /gene="flsp-12"
                /map="A3-B1"
                /number=2
     exon       1675..1926
                /gene="flsp-12"
                /map="A3-B1"
                /number=3
     intron     1927..2058
                /gene="flsp-12"
                /map="A3-B1"
                /number=3
     exon       2059..2270
                /gene="flsp-12"
                /map="A3-B1"
                /number=4
     intron     2271..2641
                /gene="flsp-12"
                /map="A3-B1"
                /number=4
     exon       2642..4006
                /gene="flsp-12"
                /map="A3-B1"
                /number=5
BASE COUNT    1061 a    1039 c    984 g    1044 t
ORIGIN   chromosome 10.

DB 55;  Score   354;  Match 76.5%;  QryMatch 17.1%;  Pred. No. 4.05e-224;
Matches  720;  Conservative   0;  Mismatches 192;  Indels  29;  Gaps 19;

Db  2640  agaaggcaaaaagtgcatccggacaacctaaaatcgcaagcctgtcaagtttgagcttt 2699
          ||||||||||||||||||||||||| || ||||||| ||||||||| |||||||||||||||
Qy   881  AGAAGGGCAAAAAGTGCATCCGTACTCCCAAAATCTCCAAGCCTACTCCAAGTTTGAGCTTT 940

Db  2700  ctggctgccacagtgtgaagacatacagggctaagttctgcgggtgtgcacagacggcc 2759
          ||||||||||||  |||||||||||| |||||||| ||| ||  || || |||||||
Qy   941  CTGGCTGCCACCAGTGAAGACATACCGAGCTAAATTCTGTGGAGTATGTACCGACGGCC 1000

Db  2760  gctgctgcacccgcacagaaccacccactctgccagtggagttcaaatgcccgatggcg 2819
          | ||||||| || ||||||||||||||||| |||||| |||||||||| |||| ||||
Qy  1001  GATGCTGCACCCCCACAGAACCACCACCCTGCCGGTGAGTTCAAGTCCCCTGACGGCG 1060

Db  2820  agatcatgaaaagaatatgatgttcatcaagacctgtgctgccattacaactgtcctg 2879
          || ||||||| ||||| ||||||||||||||||||||||||||||||||||||||||
Qy  1061  AGTTCATGAAGAGACAACATGATGTTCATCAAGACCTGTGCCTGCCATTACAACTGTCCCG 1120

Db  2880  gggacaatgacatctttggatcctgtactacacggagagatgtacggagacatggcgtaa 2939
          ||||||||||||||||||| || |||||||||||||||||||||||||||||||| | ||
Qy  1121  GAGACAATGACATCTTTGAATCCTGTACTACACAGGAAGATGTACGGAGACATGGCATGAA 1180

Db  2940  gccaggaagtaaggagacgaactcattagaactaaacttgaactgagttgcatctcatt 2999
          ||||| ||| ||| |||||||||||||||||| ||||||||||||| ||||||||||||||
Qy  1181  GCCAGAGAGTGAGAGACATTAACTCATTAGACTTGAACTTGAACTGAGTTCAAGTCTCATT 1240

Db  3000  ttcttctgtaaaacaattacagtagcacat-taatttaaatctgtgttttaactaccg 3058
          || || ||||||| ||| |||||||| | ||||||||||||||||||| ||| |||
Qy  1241  TT--TCCGTAAAAATGATTTCAGTAGCACAAGTTATTTAAATCTGTTTTCTAACTG--G 1296

Db  3059  tgggaggaactatcccaccaaagtgagaacgttatgtcatggccatacaagtagtctgtc 3118
          ||| | | |||||| || ||||| | ||| |||| ||| |||||| || ||||| |||| ||
Qy  1297  -GGGAAAGAT-TCCCACCCAATTCAAAACATTGTCCCATGTCAA-ACAAATAGTCTATC 1353
```

Page 20:

```
Db  3119  aacctcagacactgtttcgagacagttacacttgacagttgt--t-cattagccgcaca 3175
          || ||||||||||||||||| ||| ||| ||||||||| | |||||| |||
Qy  1354  TTCCCAGACACTGGTTGAAGAGATTTAAGACTTAAGACTTGCAACAGTGGAACTACATTAGTACACA 1413

Db  3176  gtgccagaaacgcacactgaggtgagtcctgaacagtggaatgccaggagaagaaa 3235
          | |||||||| |  |||||||  |||| || ||||  |||| ||| |
Qy  1414  GCACCAGAATGTATTATTAAGGTGTGGCTTTAGGAGCAGTGG-GAGGGTACCGCCCG--- 1469

Db  3236  gacaggtactagctgaggttattttaaaagcacagcagtgtgccacttttggagtgtgaac 3295
          || ||| |||| ||||||| || | || ||  || |||| ||||||||||| ||||||||
Qy  1470  GTTAG-TATCATCAGATCGACTCTTTATACG-AGTAATATGCCTGCTATTTGAAGTGTAAT 1527

Db  3296  cgggggaggaaattatagcatgcttgcagacagacctg---cctcagcggagctgagca 3352
          |||||||||||||| ||| || ||| |||||  ||||   || | || |||||||||
Qy  1528  TGAGAAGGAAAATTTTAGCCTGCTGCCTGACCTGCCTGAGCCCCAGTGACAGCCTAGGAT 1587

Db  3353  -tgtgtcctcca-ctag-atgaggctgagtccagctgttcttttaaga-acagcagtttca 3408
          | | ||||| ||| ||||| | ||| |  |||||||||  | | |||
Qy  1588  GTGCATTCTCCAGCCATCAAGACACTGAGTCAAGTTGTTCCTTAAGTCAGACAGCAGAC 1647

Db  3409  gcctctgac-cattctgattccagtgacacttgtcaggagtcagagccctgctgttaga 3467
          | || |||||||||||||| ||||||  | | | || || |||||| ||||||
Qy  1648  TCAGCTCTGACATTCTGATTCGAATGACACTGTTCAGGAATCGGAATCCTGTTCGATTAGA 1707

Db  3468  ctggacacgcttgtggcaagtaagttgcctgtaacaagccagagatttt----attgatat 3523
          ||||||||||||||||||| || ||||||||||||||||||||||||   ||| ||||
Qy  1708  CTGGACAAGCTTGTGGCAAGTGAAGTGATTTGCCTGTAACAAGCCAGATTTTTTAAAATTTATAT 1767

Db  3524  tgtaaatattgtggatataaataatgtatatatatatatttgta 3564
          |||||||||||||| | ||||||| || ||||| | |||||||||
Qy  1768  TGTAAATATTGTGTGTGTGTGTGTGTGTGTATATATATA 1808

RESULT        7
LOCUS         GGNOVMRNA       1976 bp    RNA           VRT        08-JAN-1992
DEFINITION    G.gallus nov mRNA.
ACCESSION     X59284
KEYWORDS      NOV gene.
SOURCE        chicken.
  ORGANISM    Gallus gallus
              Eukaryotae; mitochondrial eukaryotes; Metazoa/Eumycota group;
              Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata;
              Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Choanata;
              Tetrapoda; Amniota; Reptilia; Sauria; Archosauria; Aves;
              Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
REFERENCE     1  (bases 1 to 1976)
  AUTHORS     Perbal,B.
  TITLE       Direct Submission
  JOURNAL     Submitted (19-APR-1991) to the EMBL/GenBank/DDBJ databases. B.
              Perbal, Institut Curie, Centre Universitaire Batiment 110, 91405 -
              Orsay, Cedex, FRANCE
REFERENCE     2  (bases 1 to 1976)
  AUTHORS     Joliot,V., Martinerie,C., Dambrine,G., Plassiart,G., Brisac,M.,
              Crochet,J. and Perbal,B.
  TITLE       Proviral rearrangements and overexpression of a new cellular gene
              (nov) in myeloblastosis-associated virus type 1-induced
              nephroblastomas
  JOURNAL     Mol. Cell. Biol. 12 (1), 10-21 (1992)
  MEDLINE     92107157
COMMENT       NCBI gi: 63702
FEATURES      Location/Qualifiers
     source   1..1976
              /organism="Gallus gallus"
```

```
Db   632 ttgaacagcaacaggaggaagtgcttgttccaaagctgtggaatgggcgtttctacccc  691
             | ||||||| |||| |||| || ||| ||||||| |  || ||||| ||| | ||| 
Qy   728 TGGTCCAGACCACAGAGTGGAGCCCCTGTTCCAGACCTGTGGATGGGCATCTCCACCC   787

Db   692 gtgttaccaacagaaatcagcagtgtgagatggtgaagcagcacgacctttgcatgatga  751
             | ||||||||   |  | |||| | || | |||  || ||||| | | ||| |||||| 
Qy   788 GGGTTACCAATGACAACGCCCTCCTGCAGGCTAGAGACGCAGCAGCCGCCTGTGCATGGTCA  847

Db   752 gaccttgaaacgaaacgaagagccatctga---taagaaaggaaaaaatgtatccaaacaa  808
             | ||||| |  ||| |||| || ||  |      |||| ||||||   |   |  |  | 
Qy   848 GGGCCTTCCGAAGCTGACCTGGACAGAGACATTAAGAGGGCCAAAAACTGCATCCGTACTC   907

Db   809 agaaatccatgaaagctgttcgttttgaatacaagaacgtgcaccagtgtgcagacttaca  868
            | | |  ||   ||   | |  | || |||| | ||| |  |||||    |        | 
Qy   908 CCAAAATCCAACCCATCAAGTTCAGCTTTCTGCCTGCACCAGCATGAAGAGACATACC    967

Db   869 aacctcgttactggcctcgcaatgatgggcgatgctgtacccccacacaacaccaaaaa  928
            | |  |  ||| || ||| ||  |||| |  | || | ||| ||| |  |  |  | ||| 
Qy   968 GAGCTAAATTCTGTGGAGTATGTACCCACCGCCCCATCCTCACCCCCACCAGAGACCACCA  1027

Db   929 cgattccaagttgagttccgtgtcctcagggcaaattcctaaaaagccaatgatgttga  988
            |  | |  |||  | ||||| |  | |||  ||||  | |  |||||| ||| ||| |  | 
Qy  1028 CCCTGCCCGTGGAGTTCAAGTGCCCTCACCGCCGAGGTCATGAAGAACATGATGTTCA   1087

Db   989 tcaataccctgtctcgtcatggtaactgtcc 1019
            ||  | || ||| ||| |||  | |  | || 
Qy  1088 TCAAGACCTGTGCCTGCCATTACAACTGCC  1118
```

RESULT   8
LOCUS   CU13063   1927 bp   mRNA   VRT   28-AUG-1994
DEFINITION   Coturnix japonica cellular proto-oncogene Nov (nov) mRNA,
             complete cds.
ACCESSION   U13063
KEYWORDS   .
SOURCE   Coturnix japonica.
ORGANISM   Coturnix japonica
             Eukaryota; Hyperchondria; Eukaryote crown group; Metazoa/Eumycota
             group; Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia;
             Chordata; Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii;
             Choanata; Tetrapoda; Amniota; Reptilia; Sauria; Archosauria; Aves;
             Neognathae; Galliformes; Phasianidae; Phasianinae; Coturnix.
REFERENCE   1 (bases 1 to 1927)
AUTHORS   Weiskirchen,R. and Bister,K.
TITLE   Suppression in transformed avian fibroblasts of the nov
             protooncogene
JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 1927)
AUTHORS   Bister,K.
TITLE   Direct Submission
JOURNAL   Submitted (08-AUG-1994) Bister K., University of Innsbruck,
             Institute of Biochemistry, Peter-Mayr-Strasse 1a, Innsbruck,
             Austria, A-6020
COMMENT   NCBI gi: 532696
FEATURES   Location/Qualifiers
     source   1..1927
                     /clone="w23o"
                     /clone_lib="QEF cDNA library of R. Weiskirchen and K.
                     Bister"
                     /organism="Coturnix japonica"
                     /cell_type="fibroblast"
                     /dev_stage="embryo"
     CDS   14..1075
                     /gene="nov"

```
                     /strain="Brown Leghorn"
                     /dev_stage="13 day embryo"
                     /cell_type="fibroblast"
                     /clone_lib="cDNA-ELB"
                     /clone="pC1K"
     exon            1..92
     CDS             25..1080
                     /gene="nov"
                     /note="NCBI gi: 63703"
                     /codon_start=1
                     /translation="METGGGQGLPVLILLLLLLRPCEVSGREAACPRPCGGRCPAEPP
                     RCAPGVPAVLDGCGCCLVCARORGESCSPLLPCDESGGLYCDRGPEDGGGAGICMVLE
                     GDNCVFDGMIYRNGETFQPSCKYQCTCRDGQIGCLIPRCNLGLLLPGPDCPFPRKIEVP
                     GECCEKWVCDPRDEVLLGGFAMAAYRQEATLGIDVSDSSANCIEQTTEWSACSKSCGM
                     GFSTRVTNRNQQCEMVKQTRLCMMRPCENEEPSDKKGKKCIQTKKSMKAVRFEYKNCT
                     SVQTYKPRYCGLCNDGRCCTPHNTKTIQVEFRCPQGKFIKRPMMLINTCVCHGNCPQS
                     NNAFFQPLDPMSSEAKI"
     exon            93..321
                     /gene="nov"
     exon            322..573
                     /gene="nov"
     exon            574..882
                     /gene="nov"
     exon            883..1975
     polyA_signal    1815..1820
     polyA_signal    1933..1938
BASE COUNT      522 a    452 c    486 g    516 t
ORIGIN
```

DB 74; Score   264; Match 66.0%; QryMatch 12.7%; Pred. No. 9.46e-159;
Matches 575; Conservative   0; Mismatches 293; Indels   3; Gaps   1;

```
Db   152 cgccgcgctgcgcccccggagagtgccgccgtcggctgcggctgcctgttgtt  211
            ||||| || | ||  ||||   ||| ||||||||||||||||||| ||| || 
Qy   248 CGCCGCCCTGCCCGCGCCGGGCGGGCCTCGTGCTGGACGGCGTCGCGGTGCTGCCGCCTCT  307

Db   212 gggccggcaagcgcgcgcgaagctgctccctcctgcctgtgcgacgacgagcggcggcc  271
            || ||  ||||  | || ||||   |||||  | ||||||| |||   ||||||||||||| 
Qy   308 GCGCCAAGCAGCTCGGGCGAGCGTGCACCCAGCCCGACCCCCTGCGACCCCCACAAGGGCC  367

Db   272 tctactgcgacgcggccccgagacgggacggcggcgccggcgcatctgcatggtgctggaag  331
            ||| || || | ||||| |  || |||| || |||||  | || ||  || ||| || || 
Qy   368 TCTTCTGTGACTTCGGCCTCCCCGGCCAACCGCAAGATCGGCGTGCGACCCGCAAAGATG  427

Db   332 gggacaactgcgtgttcgatggatgatttaccgccaacggggagacgttccagcccagct  391
            ||| ||  |  |  | |||| | ||    |||||  | |  ||||| | |||||| | |||| 
Qy   428 GTGCTCCCTCGCATCTTCGGTGGTACGGTGTACCCCAGCGGAGAGTCCTTCCAGACGAGCT  487

Db   392 gcaagtaccagtgcacctgcggagacgggcagatcggggcgctgccccgctgcaacctgg  451
            ||||||||||||||||||||| | ||||| | |  || ||  | ||||| | ||  |||| 
Qy   488 GCAAGTACCAGTGCACCTGCCTGACCGGGCGGTGGGCTCATGCCCCTGCAGCATGG  547

Db   452 gctgctgctcccggcccccgactgccctttcccgcggaagatccgagtcccgcccgaggagt  511
            | | |  | |||  |||| ||||||| |||||||| || || | | | ||| ||||||||| 
Qy   548 ACGTTCGTTCGTCCTGCCCAGCCCTGACTGCCCCCTTCCCGAGGAGGTCAAGCTCCCCGGGAAAT  607

Db   512 gctgcgagaagtgggtgtgcgacccccaggggatgaagtgctcctgggagcttgctatgg  571
            ||||||  |   ||||| || ||   ||||| |||||| ||||  ||| | | |||| |  
Qy   608 GCTGCGGAGGAGCTGGGTCGTGTGACGACGCCCAAGACCAAACCGTGGTTGGGCCTGCCCTCG  667

Db   572 ctgcatacagacaggaggccacacttggatagacgtgtctgattcaagtgccaattgta  631
            ||||||||| || ||||| ||| ||||||  || ||| ||||| |||| || |  |||  
Qy   668 CGGCTTACCCGACTGGAAGACACGTTTGGCCCAGACCCAACTATGATTAGAGCCAACTGCC  727
```

```
                /note="NCBI gi: 532697"
                /codon_start=1
                /product="cellular proto-oncogene protein Nov"
                /translation="MEPGGGHSLPVLLLLLLLLLLRPSEVNGREAPCRPRCCGRCPAE
                PPRCARPGVPAVLDGCGCCLVCARQRGESCSPLLPCDESGGLYCDRGPEDGGGTGICMV
                LEGDNCVFDGMIYRNGETFQPSCKYQCTCRDGQIGCLPRCNLGLLLLPGDCPFPRKIE
                VPGECCEEKWVCEPRDEVLLGGFAMAYRQEATLGIDVSDSSANCIEQTEWSACSRSC
                GMGFSTRVTNRNQQCEMVKQTRICMMRPCENEEPSDKKGKKCIRTKKSMKAVRFEYKN
                CTSVQYTKPRYCGLCNDGRCCTPHNTKTIQVEERCPQGKFIKKPMMLINTCVCHGNCP
                QSNNAFFQPLDPMSSEAKI"

BASE COUNT      473 a     450 c     489 g     515 t
ORIGIN

DB 73;  Score   569;  Match 65.3%;  QryMatch 12.1%;  Pred. No. 4.22e-150;
Matches  252;  Conservative   0;  Mismatches 299;  Indels   3;  Gaps   1;

Db  147  cgccgcgctggccgccccgggggtgccgcagtgctgacggctgcggctgttgctgttgt  206
         ||||| ||||| |  |||||||  |  || ||| ||| ||| |||||||| | ||||
Qy  248  CGCCGGCCTGCCGGGGCGCGGCCGCGTCGCCGGCGCCGGCCGGTGCCTCGTGCCGGCTCT  307

Db  207  gggccgccgcagcgcggagagagctgaccctctgctcctgcgacgagagcgggcggc  266
         ||||| || || ||  ||    ||   || |||| ||||||      ||||||
Qy  308  GCGCCAGCAGCTGGGCGGCCAGCGTGCCACCGGACCCCTGCACCCGCACCAAGGGCC  367

Db  267  tctactgcgaccggcccgaggacggcggaggcaccggcatctgatggtgctgaag  326
         |||| ||||| |||||||||||  |||| || |||||  ||| || ||||||
Qy  368  TCTTCTGTGACTTCGGCATCGGCTGCCTCCCGCGCTGTAACCTCGGCCTGCTGCTGCTCCCGGGCGACTGCCCTTTCCCAAAGATG  427

Db  327  gggacaactggctgtcgatggatgattaccgcaacggaggagacgttccagccagct  386
         ||||| | ||| ||| |||  || || | ||||| ||||||| || ||  || ||
Qy  428  GTGCTCCCTGCATCTTCGGTGTGACGGTGTACCGGTGTACCCCAGCGGCGGAGACGTCCTTCCAGACCAGCT  487

Db  387  gcaagtaccaagtgcacctgccggcagacggcgagatcggctgcctgccccgtgcaacctcg  446
         ||||| ||||||||| |||||| |||  || ||| |||| |  ||| |   | ||||
Qy  488  GCAAGTACCAGTGCACGTGCCTGCGACGGGGTGGCTGCGGGCGCCTGCCCCTGTGCCAGCATGG  547

Db  447  gcctgctgtcccgggtccgactgccccttcccggaagaatcgaagtccccgggagagt  506
         ||||| || |||   ||| || ||| |  ||| ||| || || ||||| | |
Qy  548  AGTTCGTCTGCCCAGCCTCGACTGCCCCTTCCCCTCCCGAGGAGTCAAGCTGCCCGGGAAAT  607

Db  507  gctgcgagaagtggggtctgtgagccaggatgaagtgctgctcctgggaggctttgctatgg  566
         || || | |||| |||||||||  |||||| || || || ||    ||||  ||| |||
Qy  608  GCTCCGAGGACTGGGTGTGTGACAGCAAGGACGCCCTGGAGGAGGAGCTGGAGGGCTTCTTCTCG  667

Db  567  ccgcatacagacaggaggccacactgggatagatgtgtctgattcaagtgccaattgta  626
         || ||||| || |||  || ||| |  ||| || || |  || |  || ||   || |
Qy  668  CGGCTTACCGACTGCGAAGACACAGTTTGGCTGCCAACTATGATTAGAGCCAACTGCC  727

Db  627  tcgagcagacaacagaaatgagacgcttgttcccagaagctgtgaatggcgtttctaccc  686
         ||||| |||| || | ||| || ||||| || |  ||| || ||||||| ||  | ||
Qy  728  TGGTTCCAGACCACCAGAGTGGACACCCGTGCCAGACCTGTTGCCAAGACCTGTGGGACATGGGCATCTCCACCC  787

Db  687  gtgttaccaacagaaaatcagcagtggaggtggtaagacagacgacttgcatgatga  746
         ||||||||||| | ||||| |||  |||||| || |||| |||| |||||||||||
Qy  788  GGGTTACCAACGATAACCCCTGCGACCTCATCAAGGAGACGAGGATCTGCGAGGTCA  847

Db  747  gacctgtgaaaatgaagagctgatctga---taagaagaaggaaaaaatgtatccgaacaa  803
         || | || |||  | || | |    |       |  || || |||   || | | |
Qy  848  GGCCTTGCGGAGCTGACCTGGAAGACAGAGACCATTAAGAGGGCAAAAAGTCGATCCGTACTC  907

Db  804  agagtcagatgaaagctgtcgtttgaataacaagaacttgcaccagtgtcagacttaca  863
         |  ||  || ||||| |||  |||  |  |   |||| | || ||||| |||  |||
Qy  908  CCAAAATCTCCAAGCCTATCAGTTTGAGCTTTGAGCTTGGCGTTCTGCTGCACCATGAAGACATACC  967
```

```
Db  864  aaccgcgttactggcctctggcctctgcaatgatgggcgatgctgtaccccaccaacaacaaaa  923
         | |||||||  |||| | | ||||| || |||||||  |||||||||| |||| ||||
Qy  968  GAGCTAAATTCTGTGGAGTATGTACCCACGGCCCATGCTGCCATCCCCCCAGAACCACCA  1027

Db  924  cgattcaagttggttccgctgtcctcaggcaaatcctgaaaagccaatgatgttga  983
         || |  || |||||| || ||||| | |||  |||| || |||||||||  |||||||
Qy  1028 CCCTGCCGGTGGAGTTCAAGTGCCTGACCGCGAGGTCATGAAGAACATGATGTTCA  1087

Db  984  tcaatacctgtctgcatggtaactgtcc  1014
         |||| |||||| ||| ||  ||||||
Qy  1088 TCAAGACCTGTGCCTGCCTGCCATTACAACTGTCC  1118

RESULT  9
LOCUS       CHKCEF       1805 bp    mRNA            VRT       15-SEP-1989
DEFINITION  Chicken CEF-10 protein mRNA, complete cds.
ACCESSION   J04496
KEYWORDS    CEF-10 protein; TPA-repressible protein;
            phorbol ester repressible protein; pp60 v-src inducible protein;
            v-srg inducible protein.
SOURCE      Chicken embryo fibroblasts, cDNA to mRNA, infected with Rous
            sarcoma virus mutant tsNY72-4, clone CEF-10.
ORGANISM    Gallus gallus
            Eukaryota; Animalia; Chordata; Vertebrata; Aves; Neornithes;
            Neognathae; Galliformes; Phasianidae.
REFERENCE   1 (bases 1 to 1805)
AUTHORS     Simmons,D.L., Levy,D.B., Yannoni,Y. and Erikson,R.L.
TITLE       Identification of a phorbol ester-repressible v-src-inducible gene
JOURNAL     Proc. Natl. Acad. Sci. U.S.A. 86, 1178-1182 (1989)
MEDLINE     89145206
COMMENT     Draft entry and computer-readable sequence for [1] kindly submitted
            by D.L.Simmons, 03-MAR-1989.

FEATURES             Location/Qualifiers
     source          1..1805
                     /organism="Gallus gallus"
     mRNA            <1..1805
                     /note="CEF-10 mRNA"
     sig_peptide     54..119
                     /note="CEF-10 protein signal peptide"
                     /codon_start=1
     CDS             54..1181
                     /note="CEF-10 protein precursor; NCBI gi: 211436"
                     /codon_start=1
                     /translation="MCSAGARPALAAALLCLARLALGSPCPAVCCQCPAVCQCPAAAPQCAPGVG
                     LVPDGCGCCCKVCAKQLNEDCSRTQPCDHTKGLECNFGASPAATNGICRAQSEGRPCEY
                     NSKIYQNGESFQPNCKHQCTCIDGAVGCIPLCPQELSLPNLGCPSPRLVKVPGQCCEE
                     WVCDESKDALEEELEGFFSKEFGLDASEGELTRNNELIAIVKGGLKMLPVFGSEPQSRA
                     FENPKCIVQTTSWSQCSKTCGTGISTRVTNDNPDCKLIKETRICEVRPCGPSYASLK
                     KGKKCTKTKKSPSPVRFTYAGCSSVKKYRRYCGSCVDGRCCTPQQTRTVKIRFRCDD
                     GETFTKSVMIQSCRCNYNCPHANEAYPFYRLVNDIHKFRD"
     mat_peptide     120..1178
                     /note="CEF-10 protein"
                     /codon_start=1

BASE COUNT      417 a     492 c     459 g     437 t
ORIGIN       Unreported.

DB 73;  Score   127;  Match 67.8%;  QryMatch 6.1%;  Pred. No. 8.80e-62;
Matches  270;  Conservative   0;  Mismatches 125;  Indels   3;  Gaps   2;

Db  159  gcgccgcagtgcgccccggcgccgccggtgggcgtgctgggcctgctgccaaggtc  218
         |||||||  ||| |||||||| ||||||||||||| | |||||||||||||||
Qy  247  GCGCCGCCGTCGCCGGCGCCGGCCGGTGCCTCGTGCCGACGGCGTCGCCGGCGTC  306
```

```
Db  219  tgcgccaagcagctgaacgaggactgcagccgacgcagccctgcgaccaccaagggg  278
         |||||||||||  ||||   ||  | | |||||  ||  | | |||||||||  ||||||
Qy  307  TGCCCCAGCAGCTGGGCGAGCTGTGCACCGAGCAGCCGCAGCCCTGCGACCCGCCACAAGGGC  366

Db  279  ctggagtgcaacttcggcgccagccccgcgcgccaccaacggcatctgcagagacagtct  338
         ||   |||||||| || ||   |  |    || |  ||||||| |  |||  ||  || ||
Qy  367  CTCTTCTGTGACTTCGGCTCCCCGGCCAACCGCAAGATCGGCGTGTCCACCGC-CAAA--  423

Db  339  gaggggaccatgcaatacaactccaaaatctaccagaacagcgaaagcttccagacccc  398
         || ||  ||                |   |||  |    | || |||| |||||  ||
Qy  424  GATGGTCTGCCCCGCATCTTGGTGTCGCATCTTCGGTGGTACCGTGTACCGCCGGAGAGTCC  483

Db  399  aactgaagcaacagtgtactgcatagatggacgtgtgggctgcatcccgtctgccg  458
         | ||||||| ||| || ||  | |  ||  |||||| ||||| ||| |||  || |
Qy  484  AGTTGCAAGTACCAGTCCAGGTGCGACGGGCGTGGCGTGGGCTGCATGTCCCCTGTCCAGC  543

Db  459  caggagctctcctccccaacctgggtgccccagcccaggctggtcaaagtgcctggg  518
         |||||| |||| |||   || || | |||||  |||  || |||| ||||||| |||
Qy  544  ATGGACGTTCGTTGCCCAGCCCTGACTGCCCCTGACGGAGGTCAAGTGCCCCGGG  603

Db  519  cagtgctgcgaggagtgggtctgcgatgagagcaaga  556
         ||||| || | | ||| |||| || ||||||| |||
Qy  604  AAATGCTGCGAGGAGTGGGTGTGTGACGAGCCCAAGGA  641
```

RESULT 10
LOCUS   MUSCYR61A   2018 bp   mRNA   ROD   12-JUN-1993
DEFINITION Mouse Cyr61 mRNA, complete cds.
ACCESSION   M32490
KEYWORDS   cyr61 product.
SOURCE   Mouse (strain BALB/c) fibroblast cell line 3T3, cDNA to mRNA.
ORGANISM   Mus musculus
            Eukaryota; Animalia; Chordata; Vertebrata; Mammalia; Theria;
            Eutheria; Rodentia; Myomorpha; Muridae; Murinae.
REFERENCE   1 (bases 1 to 2018)
AUTHORS   O'Brien,T.P., Yang,G.P., Sanders,L. and Lau,L.F.
TITLE   Expression of cyr61, a growth factor-inducible immediate-early gene
JOURNAL   Mol. Cell. Biol. 10, 3569-3577 (1990)
MEDLINE   90287146
COMMENT   Draft entry and computer-readable sequence [1] kindly submitted by
            L.F.Lau, 02-MAR-1990, for release after publication.

            NCBI gi: 192909
FEATURES   Location/Qualifiers
source   1..2018
            /organism="Mus musculus"
mRNA   <1..2018
            /note="Cyr61 mRNA"
CDS   190..1329
            /note="Cyr61 product; NCBI gi: 309206"
            /codon_start=1
            /translation="MSSSTFRTLAVAVTLLHLTRLALSTCPAACHCPLEAPKCAPGVG
            LVRDGCGCCVCAKQLNEDCSKTQPCDHTKGLECNFGASSTALKGICRAQSEGRPCEY
            NSRIYQNGESFQPNCKHQCTCIDGAVGCIPLCPQELSLPNLGCPNRLVKVSGQCCEE
            WVCDEDSIKDSLDDQDDLLGLDASEVELTRNNELIAICKGSSLKRLPVFGTEPRVLFN
            PLHAHGQKCIVQTSWSQCSKSCGTGISTRVTNDNPECRLVKETRICEVRPCGQPVYS
            SLKKGKCSKTKKSPEPVRFTYAGCSSVKKYRPKYCGSCVDGRCCTPLQTRIVKMRFR
            CEDGEMFSKNVMMIQSCKCNYNCPHPNEASFRLYSLFNDIHKFRD"
repeat_region 1453..1551
            /note="49 bases repeat"
BASE COUNT   500 a   536 c   527 g   455 t
ORIGIN

```
DB 55;  Score  88;  Match 61.1%;  QryMatch 4.2%;  Pred. No. 1.39e-35;
Matches  243;  Conservative  0;  Mismatches 155;  Indels  0;  Gaps  0;

Db  865  aaatgcatcgttcagacacagtcttggtcccagtgctccaagagctgcggaactggcatc  924
         ||||||||| || |||||||  |||||||||||  ||||| | ||||||||||  |||||||
Qy  721  AACTGCCTGTCCAGACCACAGAGTGAGCCGCCCTGTTCCAAGACCTGTGGGATGGGCATC  780

Db  925  tccacacgagtaccaatgacacaccagagtgccgcctgtgaaagagaccggatcgt  984
         |||||||| ||||||| ||    |||| | ||||  | ||||||||||||| || ||
Qy  781  TCCACCCGGGTTACCAATGACAACGCCCTCCTGCAGGCTAGAGAAGCAGAGCCGCCTGC  840

Db  985  gaagtgcgtccttgtggacaaccagtgtacagcagctaaaaagggcaagaaatgcagc  1044
         | |||||| || ||| ||  ||||||| |||||||   |||||||| ||| |||||||
Qy  841  ATGGTCAGCCCTTGCGAAGCTGACCTGGAGAGACATTAAGAGGGCAAAAAGTGCATC  900

Db  1045 aagaccaagaaatcccagaaccagtcagattacttatgcaggatgtccagtgtcaag  1104
         || ||||||  |  ||  |||||  ||  || ||||||||| |||| || ||| |||
Qy  901  CGTACTCCAAAATCCAAGCCTATCAAGTTGAGCTTTCTGCCTGCCACCAGCATGAAG  960

Db  1105 aaatacgggcccaaatactcggctcctgcgtagatggccggtgctgcacacctctgcag  1164
         ||||||| ||||||| ||| ||  ||| |||||||| |||||||||||  |||| ||
Qy  961  ACATACCAGCTAAATTCGTGGAGTATGTACCCAGCCCGATCGTCGCACCCCCCACAGA  1020

Db  1165 accagaactgtgaagatgcggttccgatgcgaagatgcggaagatgttccaagaatgtc  1224
         |  ||||| || ||| |||  ||| || || || ||||| |||||||||||| || |||
Qy  1021 ACCACCACCGCTGCCGGTGGAGTTCAAGTGCCCTGACGGCGAGGTCATGAAGAAGACACATG  1080

Db  1225 atgatgatccagtcctgcaaatgtaactacaactgccc  1262
         ||||||| ||||||| || |||||||||| ||||||| ||
Qy  1081 ATGTTCATCAAGACCTGTGCCTGCCATTACAACTGTCC  1118
```

RESULT 11
LOCUS   HSNOVH5   1469 bp   DNA   PRI   01-DEC-1994
DEFINITION H.sapiens novH gene exon 5.
ACCESSION   X78354
KEYWORDS   novH gene.
SOURCE   human.
ORGANISM   Homo sapiens
            Eukaryota; mitochondrial eukaryotes; Metazoa/Eumycota group;
            Eukaryotae; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata;
            Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata;
            Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Choanata;
            Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates;
            Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1469)
AUTHORS   Martinerie,C., Huff,V., Joubert,I., Badzioch,M., Saunders,G.,
            Strong,L. and Perbal,B.
TITLE   Structural analysis of the human nov proto-oncogene and expression
            in Wilms tumor
JOURNAL   Oncogene 9 (9), 2729-2732 (1994)
MEDLINE   94336229
REFERENCE   2 (bases 1 to 1469)
AUTHORS   Perbal,B.
TITLE   Direct Submission
JOURNAL   Submitted (21-MAR-1994) to the EMBL/GenBank/DDBJ databases. B.
            Perbal, Institut Curie, Centre Universitaire Batiment 110, 91405
            Orsay, Cedex, FRANCE
COMMENT   NCBI gi: 587425
FEATURES   Location/Qualifiers
source   1..1469
            /organism="Homo sapiens"
            /germline
            /tissue_type="placental"
            /clone="LC110 clone 2"

```
                   /chromosome="8"
                   /map="8q24.1"
 misc_feature   158
                   /note="SA4 (splice acceptor 4)"
                   /gene="novH"
 misc_feature   159..455
                   /note="coding part of exon"
                   /label=stp
                   /usedin=X78351:HsCDS
 exon           159..>455
                   /gene="novH"
                   /number=5
                   /label=ex5
                   /usedin=X78351:HsCDS
 BASE COUNT      472 a   315 c   270 g   412 t
 ORIGIN

 DB 46;  Score    68;  Match 63.6%;  QryMatch 3.3%;  Pred. No. 9.25e-23;
 Matches 159;  Conservative   0;  Mismatches  91;  Indels   0;  Gaps   0;

 Db   157 agaaaggaaaaaagtgtctccgcaccaagaagtcactcaaagccatccacctgcagttca 216
              ||||| || |||||||| |||| ||||| || |||| | |||| |||||||||||||||
 Qy   881 AGAAGGCAAAAAGTGCATCCGTACTCCGTACTCCCAAAATCTCCAAAGCCTATCAAGTTTGAGCTTT 940

 Db   217 agaactgcaccagcctgcacacctacaagccaggttctgtggggtctgcagtgatggcc 276
              |||||||| | | || | | | |||| | | ||||  ||||||| |||| || ||| |
 Qy   941 CTGGCTCGCACCAGCATGAAGACATACCGACCTAAATTCTGTGGAGTATGTACCGACGGCC 1000

 Db   277 gctgctgcactcccacaataccaaaccatccaggcaggcagagttcagtgctcccagggc 336
              ||||| || |||||| ||| |||| || ||||||| ||| ||| ||||| |||||| |
 Qy  1001 GATGCTGCACCCCCCACAGAACCCACCACCTGCCGGTGGAGTTCAAGTGCCCTCACGGCG 1060

 Db   337 aaatagtcaagaagccagtcagtgatggtcattgggacctgcacctgtcaccaaactgtccta 396
             | | ||||| |||| | ||| |||| | ||||| ||| |||| ||| | ||||||| ||||||||
 Qy  1061 AGGTCATGAGCAAGAACATGATGTTCATCAAGACCTGCGCCATTACAACTGTCCCG 1120

 Db   397 agaacaatga 406
             |||||||
 Qy  1121 GAGACAATGA 1130

 RESULT   12
 LOCUS       HSNOVH4       435 bp    DNA                PRI       01-DEC-1994
 DEFINITION  H.sapiens novH gene exon 4.
 ACCESSION   X78353
 KEYWORDS    novH gene.
 SOURCE      human.
             Homo sapiens
 ORGANISM    Homo sapiens
             Eukaryotae; mitochondrial eukaryotes; Metazoa/Eumycota group;
             Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata;
             Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Choanata;
             Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates;
             Catarrhini; Hominidae; Homo.
 REFERENCE   1 (bases 1 to 435)
 AUTHORS     Martinerie,C., Huff,V., Joubert,I., Badzioch,M., Saunders,G.,
             Strong,L. and Perbal,B.
 TITLE       Structural analysis of the human nov proto-oncogene and expression
             in Wilms tumor
 JOURNAL     Oncogene 9 (9), 2729-2732 (1994)
 MEDLINE     94336229
 REFERENCE   2 (bases 1 to 435)
 AUTHORS     Perbal,B.
 TITLE       Direct Submission
 JOURNAL     Submitted (21-MAR-1994) to the EMBL/GenBank/DDBJ databases. B.
```

```
 COMMENT     Perbal, Institut Curie, Centre Universitaire Batiment 110, 91405
             Orsay, Cedex, FRANCE
             NCBI gi: 587424
 FEATURES    Location/Qualifiers
 source      1..435
                   /organism="Homo sapiens"
                   /germline
                   /tissue_type="placental"
                   /clone="LC78 clone 5"
                   /chromosome="8"
                   /map="8q24.1"
 misc_feature   105
                   /gene="novH"
                   /note="SA3 (splice acceptor 3)"
 exon           105..319
                   /gene="novH"
                   /number=4
                   /label=ex4
                   /usedin=X78351:HsCDS
                   319
 misc_feature      /gene="novH"
                   /note="SD4 (splice donor 4)"
 BASE COUNT      120 a   106 c   114 g    95 t
 ORIGIN

 DB 46;  Score    65;  Match 67.2%;  QryMatch 3.1%;  Pred. No. 6.81e-21;
 Matches 127;  Conservative   0;  Mismatches   62;  Indels   0;  Gaps   0;

 Db   104 gcttacaggccagaagccacctaggagtagaagtctctgactcaagtgtcaactgcatt 163
              |||| || |||||||| ||| ||| || |||||| || | ||| |||||||||| ||
 Qy   670 GCTTACCGACTGGAAGACACGTTTGGCCCAGACCCAACTATGATTAGAGCCAACTGCCTG 729

 Db   164 gaacagacacagagtggacagcatgctccaagagtcgtggtaggggttctccacccgg 223
              ||||||||||||| |||| |||| ||||||| ||||||||||| ||||||||||||
 Qy   730 GTCCAGACCACCAGAGTCGAGCGCCTGTTCCAAGACCCTGCGGATGGGCATCTCCACCCGG 789

 Db   224 gtcaccaataggaaccgtcaatgtgagatgctgaaacagactggctctgcatggtgcgg 283
             || ||| ||| || ||| || ||||||| || | ||| || |||||| || |||||| ||
 Qy   790 GTTACCAATGACAACACCGCCTCCGAGCGTAGAGAAGCAGAGCCCCCTGTTGCGATGGTCAGG 849

 Db   284 ccctgtgaa 292
             || || |||
 Qy   850 CCTTCCGAA 858

 RESULT   13
 LOCUS       HSNOVH3       452 bp    DNA                PRI       01-DEC-1994
 DEFINITION  H.sapiens novH gene exon 3.
 ACCESSION   X78352
 KEYWORDS    novH gene.
 SOURCE      human.
             Homo sapiens
 ORGANISM    Homo sapiens
             Eukaryotae; mitochondrial eukaryotes; Metazoa/Eumycota group;
             Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata;
             Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Choanata;
             Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates;
             Catarrhini; Hominidae; Homo.
 REFERENCE   1 (bases 1 to 452)
 AUTHORS     Martinerie,C., Huff,V., Joubert,I., Badzioch,M., Saunders,G.,
             Strong,L. and Perbal,B.
 TITLE       Structural analysis of the human nov proto-oncogene and expression
             in Wilms tumor
 JOURNAL     Oncogene 9 (9), 2729-2732 (1994)
 MEDLINE     94336229
```

```
JOURNAL     Submitted (14-NOV-1990) to the EMBL/GenBank/DDBJ databases. Lau
            L.F., University of Illinois College of Medicine, Dept. of Genetics
            (M/C 669), 808 South Wood Street, Chicago, IL 60611, USA
REFERENCE   2 (bases 1 to 5196)
AUTHORS     Latinkic,B.V., O'Brien,T.P. and Lau,L.F.
TITLE       Promoter function and structure of the growth factor-inducible
            immediate early gene cyr61
JOURNAL     Nucleic Acids Res. 19 (12), 3261-3267 (1991)
MEDLINE     9128803
COMMENT     Cyr61 is an immediate early gene that is transcriptionally
            activated in T3T fibroblasts by serum, platelet-derived growth
            factor and the tumour protein TPA.
            See M32490 for mRNA sequence.

            NCBI gi: 50632
FEATURES            Location/Qualifiers
     source         1..5196
                    /organism="Mus musculus"
                    /strain="AJ"
                    /dev_stage="embryo"
                    /cell_type="fibroblast"
                    /clone_lib="AJ strain genomic lambda"
     mRNA           join(2063..2314,2610..2823,3168..3512,3634..3848,
                    4079..>4381)
                    /gene="CYR61"
     exon           2063..2314
                    /gene="CYR61"
                    /number=1
     CDS            join(2252..2314,2610..2823,3168..3512,3634..3848,
                    4079..4381)
                    /gene="CYR61"
                    /note="NCBI gi: 50633"
                    /codon_start=1
                    /translation="MSSSTFRTLAVAVTLLHLTRLALSTCPAACHCPLEAPKCAPCVG
                    LVRDGCGCCKVCAKQLNEDCSKTQPCDHTKGLECNFGASSTALKGICRAQSEGRPCEY
                    NSRIYQNGESFQPNCKHQCTIDGAVGCIPLCPQELSLPNLGCPNRPLVKVSGQCCEE
                    WVCDEDSIKDSLDDQDDLLGLDASEVLTRNNELIAIGKGSSLKRLPVFGTEPRVLFN
                    PLHAHGQKCIVQTSWSQCSKSCGTGISTRVNDNPECRLVKETRICEVRPCGQPVYS
                    SLKKGKKCSKTKKSPEPVRFTYAGCSSVKKYRPKYCGSCVDGRCCTPLQTRTVKMRFR
                    CEDGEMFSKNVMMIQSCKCNYNCPHPNEASFRIYSLFNDIHKFRD"
     intron         2315..2609
                    /number=1
     exon           2610..2823
                    /gene="CYR61"
                    /number=2
     intron         2824..3167
                    /number=2
     exon           3168..3512
                    /gene="CYR61"
                    /number=3
     intron         3513..3633
                    /number=3
     exon           3634..3848
                    /gene="CYR61"
                    /number=4
     intron         3849..4078
                    /number=4
     exon           4079..>4381
                    /gene="CYR61"
                    /number=5
     polyA_signal   5049..5054
BASE COUNT      1365 a   1355 c   1232 g   1244 t
ORIGIN

DB 53;  Score    60;  Match 67.0%;  QryMatch 2.9%;  Pred. No. 7.98e-18;
```

```
REFERENCE   2 (bases 1 to 452)
AUTHORS     Perbal,B.
TITLE       Direct Submission
JOURNAL     Submitted (21-MAR-1994) to the EMBL/GenBank/DDBJ databases. B.
            Perbal, Institut Curie, Centre Universitaire Batiment 110, 91405
            Orsay, Cedex, FRANCE
COMMENT     NCBI gi: 587423
FEATURES            Location/Qualifiers
     source         1..452
                    /organism="Homo sapiens"
                    /germline
                    /tissue_type="placental"
                    /clone="LC48 clone 5"
                    /chromosome="8"
                    /map="8q24.1"
     misc_feature   117
                    /gene="novH"
                    /note="SA2 (splice acceptor 2)"
     exon           117..368
                    /gene="novH"
                    /number=3
                    /label=ex3
                    /usedin=X78351:HsCDS
     misc_feature   368
                    /gene="novH"
                    /note="SD3 (splice donor 3)"
BASE COUNT      111 a    96 c    117 g    128 t
ORIGIN

DB 46;  Score    62;  Match 65.0%;  QryMatch 3.0%;  Pred. No. 4.80e-19;
Matches  134;  Conservative   0;  Mismatches  72;  Indels   0;  Gaps   0;

Db   121 agaggagataactgtgtgttcgatgggtcatctaccgcagtggagagaaatttcagcc 180
             ||| ||||| | | |||| ||| | ||||||||| | |||||||| |||||| || |||
Qry  423 AGATGGTGCTCCCTGCATCTTCGGTGGTACCGGTGTACCGCAGCCGAGTCCTTCCAGAG 482

Db   181 aagctgacaaattccagtgacacctgcagagatgggcagattggctgtgtgcccgctgtca 240
             |||||||| | ||||||| | ||||||| || ||| | | |||| ||||| |||||| ||
Qry  483 CAGCTGCGAAGTACCAGTGCACGTGCCTCGGACGGGCCGTGGGCTGCATGCCCGTGTCCAG 542

Db   241 gctggatgtgctactgctgagcctaactgcccagctccaagaaaagttgaggtgcctgg 300
             ||| ||| | | |||| ||| | |||| ||| || | |||| || ||||| ||
Qry  543 CATCGGACGTTCGTCTGCCCAGCCCTGACTGCCCTTCCCGAGGAGGTCAAGCTGCCCGG 602

Db   301 agagtgctgtgaaaagtggatctgtg 326
             ||| |||| |||| | |||||
Qry  603 GAAATGCTGCCAGGAGTCGGGTGCTG 628

RESULT   14
LOCUS       MMCYR61G     5196 bp    DNA           ROD        20-MAY-1992
DEFINITION  Mouse growth factor inducible immediate early gene cyr61.
ACCESSION   X56790
KEYWORDS    cyr61 gene; growth factor-inducible gene; immediate-early gene.
SOURCE      house mouse.
ORGANISM    Mus musculus
            Eukaryotae; mitochondrial eukaryotes; Metazoa/Eumycota group;
            Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata;
            Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Choanata;
            Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Glires; Rodentia;
            Sciurognathi; Myomorpha; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 5196)
AUTHORS     Lau,L.F.
TITLE       Direct Submission
```

```
Matches  118;  Conservative    0;  Mismatches  58;  Indels    0;  Gaps   0;

Db   3209  taccaaaacgggggaaagcttccagccaactgtaaacaccagtgacatgtattgatggc 3268
                  ||||| | |||| ||   |||||||| || |||| || |||||||||||| || ||
Qy    457  TACCCAGCGCGGAGAGTCCTTCCAGCAGCTGCAAGTACCAGTGCACGTGCCTGGACGGG 516

Db   3269  gccgtgggctgcattcctctgtgtccccaagaactgtctctcccaatctgggctgtccc 3328
                  || ||||||||||||||| |||||| |   ||| ||||  | | ||| |||
Qy    517  GCCGTGGGCTGCATGCCCCTGTGCCCCCATGCACGTTCGTCTGCCCAGCCCTGACTGCCCC 576

Db   3329  aaccccgctggtgaaagtcagcgggcagtgctgtgaagagtgggtttgtgatga 3384
                  ||| || |||| | | ||||| || | |||||||||||| ||||| |||| ||
Qy    577  TTCCCGAGGAGGTCAAGCTGCCCCGGCAAATGCTGCCAGGAGTCGGCTGTGTGACGA 632


RESULT  15
LOCUS       HSNOVH12         557 bp    DNA                     PRI        01-DEC-1994
DEFINITION  H.sapiens novH gene exons 1 and 2.
ACCESSION   X78351
KEYWORDS    novH gene.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa/Eumycota group;
            Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata;
            Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Choanata;
            Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates;
            Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 557)
AUTHORS     Martinerie,C., Huff,V., Joubert,I., Badzioch,M., Saunders,G.,
            Strong,L. and Perbal,B.
TITLE       Structural analysis of the human nov proto-oncogene and expression
            in Wilms tumor
JOURNAL     Oncogene 9 (9), 2729-2732 (1994)
MEDLINE     94336229
REFERENCE   2  (bases 1 to 557)
AUTHORS     Perbal,B.
TITLE       Direct Submission
JOURNAL     Submitted (21-MAR-1994) to the EMBL/GenBank/DDBJ databases. B.
            Perbal, Institut Curie, Centre Universitaire Batiment 110, 91405
            Orsay, Cedex, FRANCE
COMMENT     NCBI gi: 587422
FEATURES            Location/Qualifiers
     source         1..557
                    /organism="Homo sapiens"
                    /germline
                    /tissue_type="placental"
                    /clone="LC84 clone I"
                    /chromosome="8"
                    /map="8q24.1"
     CDS            join(79..162,291..516,X78352:117..368,X78353:105..319,
                    X78354:159..455)
                    /gene="novH"
                    /note="NCBI gi: 825696"
                    /codon_start=1
                    /label=HsCDS
                    /translation="MQSVQSTSFCLRKQCLCLTFLILHLLGQVAATQRCPPQCPGRCP
                    ATPPTCAPGVRAVLDGCSCCLVCARQRGESCSDLEPCDESSGLYCDRSADPSNQTGIC
                    TAVEGDNCVFDGVIYRSGEKFQPSCKFQCTCRDGQIGCVPRCQLDVLLPEPNCPAPRK
                    VEVPGECCEKWICGPDEEDSLGGLITLAAYRPEATLGVEVSDSSVNCIEQTTEWTACSK
                    SCGMGFSTRVTNNRQCEMLKQTRLCMVRPCEQPEQPTDRKGKKCLRTKKSLKAIHL
                    QFKNCTSLHTYKPRFCGVCSDGRCCTPHNTKTIQAEFQCSPQIVKKPVMVIGTCTCH
                    TNCPKNNEAFLQELELKTTRGKM"
     exon           79..162
                    /gene="novH"
```

```
                    /number=1
                    /label=ex1
     misc_feature   162
                    /gene="novH"
                    /note="SD1 (splice donor 1)"
     misc_feature   291
                    /gene="novH"
                    /note="SA1 (splice acceptor 1)"
     exon           291..516
                    /gene="novH"
                    /number=2
                    /label=ex2
     misc_feature   516
                    /gene="novH"
                    /note="SD2 (splice donor 2)"

BASE COUNT       90 a      181 c      165 g      121 t
ORIGIN

DB  46;  Score     52;  Match 65.1%;  QryMatch 2.5%;  Pred. No. 4.80e-13;
Matches  112;  Conservative    0;  Mismatches   60;  Indels    0;  Gaps   0;

Db    343  cgccgcgacctcgccgccccgggtgcgccgcgtgctgctgacggctgctcatgctgctgg 402
                  ||||||| |||||| ||| ||| ||  ||||||||||||||  ||||||| |||||
Qy    245  CGGCGCCCGCCCTCCCCGCGGCGGGCCGTAGCCTCGTGCTGCGACGGCTCCGGCTGCTGCCGCCGCG 304

Db    403  tgtgtgccgccagctggcgagagctgctcagatctggagcatgcgacgagcagtg 462
                  | ||| |||||||||   ||| |||| |||||| ||  ||||| |
Qy    305  TCTGCGCCAAGCAGCAGCTGGGCGAGCCTGCACCGAGCCGCAGCCCCCTCCGCACCCCCGCACAGG 364

Db    463  gcctctactgtgatcgcagcggggaccccagcaaacagactgcatctgcac 514
                  ||||||| ||||||| | ||    ||| | ||| | ||| |||| |||||
Qy    365  GCCTCTTCTGTGACTTCGGCTCCCCGGCACCCGCAACCCAGATCGGCGGTGTGCAC 416


Search completed: Fri Sep 22 22:17:14 1995
Job time : 1134 secs.
```

```
*********************************************************

 __  __  ____   ____   ____    ____  _   _
|  \/  ||  _ \ / ___| |  _ \  / ___|| | | |
| |\/| || |_) |\___ \ | |_) || |    | |_| |
| |  | ||  __/  ___) ||  _ < | |___ |  _  |
|_|  |_||_|    |____/ |_| \_\ \____||_| |_|   (TM)

*********************************************************
```

MPsrch_nn   n.a. - n.a. database search, using Smith-Waterman algorithm

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
             Distribution rights by IntelliGenetics, Inc.

Run on: Fri Sep 22 21:50:12 1995; MasPar time 380.24 Seconds
                                          832.372 Million cell updates/sec

Tabular output not generated.

Title: >US-08-167-628-1
Description: (1:2075) from 5408040.seq
Perfect Score: 2075
N.A. Sequence: 1 CCCGGCCGACAGCCCCCGACA..........AGGAAATGTGGTAGCTCACG 2075
Comp: GGGCCGGCTGTCGGGGCTCT..........TCCTTTACACCATCGAGTGC

Scoring table: TABLE default
               Gap 6

Nmatch STD :   Dbase 0; Query 0

Searched: 225244 seqs, 76266140 bases x 2

Database: EST
                1 EST1
                2 EST2
                3 EST3
                4 EST4
                5 EST5
                6 EST6
                7 EST7
                8 EST8
                9 EST9
               10 EST10
               11 EST11
               12 EST12
               13 EST13
               14 EST14
               15 EST15
               16 EST16
               17 EST17
               18 EST18
               19 EST19
               20 EST20
               21 EST21
               22 EST22
               23 EST23
               24 EST24
               25 EST25
               26 EST26
               27 EST27
```

```
               28  EST28
               29  EST29
               30  EST30
               31  EST31
               32  EST32
               33  EST33
               34  EST34
               35  EST35
               36  EST36
               37  EST37
               38  EST38
               39  EST39
               40  EST40
               41  EST41
               42  EST42
               43  EST43
               44  EST44
               45  EST45
               46  EST46
               47  EST47
               48  EST48
               49  EST49
               50  EST50
               51  EST51
               52  EST52
               53  EST53
               54  EST54
               55  EST55
               56  EST56
               57  EST57
               58  EST58
```

Statistics: Mean 11.925; Variance 2.639; scale 4.519

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
	1	291	14.0	506	58	T98687	ye61b01.r1 Homo sapie	0.00e+00
	2	249	12.0	264	56	T92313	yel8g08.r1 Homo sapie	0.00e+00
	3	245	11.8	434	55	T86251	yd84f01.r1 Homo sapie	0.00e+00
	4	236	11.4	345	25	R37906	yh97e04.r1 Homo sapie	0.00e+00
	5	209	10.1	420	58	T98574	ye60a12.r1 Homo sapie	2.67e-317
	6	206	9.9	376	16	R06961	yf12g05.r1 Homo sapie	9.94e-312
	7	198	9.5	357	45	T49714	ya80a10.r1 Homo sapie	6.14e-297
	8	187	9.0	331	56	T92598	ye26f01.s1 Homo sapie	1.16e-276
c	9	172	8.3	410	49	T65929	yc24e08.s1 Homo sapie	3.81e-249
c	10	169	8.1	408	45	T49715	ya80a10.s1 Homo sapie	1.16e-243
	11	167	8.0	365	46	T53387	ya88f03.r1 Homo sapie	5.21e-240
	12	157	7.6	205	46	T53524	ya96b01.r1 Homo sapie	8.41e-222
c	13	155	7.5	404	39	T27642	EST10485 Homo sapiens	3.59e-218
c	14	139	6.7	140	49	T64503	yc24e08.r1 Homo sapie	2.86e-189
c	15	124	6.0	326	3	HHEA46W	H. sapiens partial cD	1.88e-162
c	16	97	4.7	427	56	T92268	yel8g08.s1 Homo sapie	4.29e-115
c	17	94	4.5	381	28	R49034	yg65h08.s1 Homo sapie	6.30e-110
c	18	91	4.4	366	24	R37824	yh97e03.s1 Homo sapie	8.79e-105
c	19	81	3.9	376	46	T53388	ya88f03.s1 Homo sapie	7.99e-88
c	20	80	3.9	416	58	T98575	ye60a12.s1 Homo sapie	3.80e-86
c	21	75	3.6	411	58	T98637	ye61b01.s1 Homo sapie	8.22e-78

Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.

NCBI gi: 748424

FEATURES   Location/Qualifiers
source   1..506
                   /organism="Homo sapiens"
                   /clone="122185"
                   /note="human"

BASE COUNT   131 a   114 c   130 g   124 t
ORIGIN

DB 58;   Score   291; Match 96.3%; QryMatch 14.0%; Pred. No. 0.00e+00;
Matches   315; Conservative   0; Mismatches   9; Indels   3; Gaps   3;

Db   1 acaacgcctctgcgaggctagagaagcagagc-gcctgtgatggtcaggccttgcaac 59
          |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Qy   800 ACAACGCCTCTGCGAGGCTAGAGAAGCAGAGCCGCCCCTGTCGCATGGTCAGGCCTTGCCAAG 859

Db   60 ngancctggaagagaacattaagaagggcaaaaagtgcatccgtactcccaaaatctcca 119
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   860 CTGCACCTGGAAGAGAACATTAAGAAGGGCAAAAAGTGCATCCGTACTCCCAAAATCTCCA 919

Db   120 agcctatccaagttgagctttctggctgcaccagcatgaagacataccgagctaaattct 179
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   920 AGCCTATCCAAGTTGAGCTTTCTGGCTGCACCAGCATGAAGACATACCGAGCTAAATTCT 979

Db   180 gtggagtatgtaccgacgcgacggcgatgctgcaccccacagaaccaccctgccggtgg 239
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   980 GTGGAGTATGTACCGACGCGACGGCGATGCTGCACCCCCACAGAACCACCACCTGCCGGTGG 1039

Db   240 agtcaagtgccctgacggcgaggtcatgaagaagaacatgatgttcattcaagacctgt 299
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1040 AGTTCAAGTGCCCTGACGGCGAGGTCATGAAGAAGAACATGATGTTCAT-CAAGACCTGT 1098

Db   300 gcctgccattacaactgtcccgga 326
          ||||||||||||||||||||||||
Qy   1099 GCCTGCCATT-ACAACTGTCCCGAGA 1124

RESULT   2
LOCUS   T92313   264 bp   mRNA   22-MAR-1995
DEFINITION yel8g08.r1 Homo sapiens cDNA clone 118142 5' similar to gb:M92934
            CONNECTIVE TISSUE GROWTH FACTOR PRECURSOR (HUMAN);.
ACCESSION   T92313
KEYWORDS   EST.
SOURCE   human clone=118142 library=Stratagene lung (#937210)
            vector=pBluescript SK- host=SOLR cells (kanamycin resistant)
            primer=M13RP1 Rsite1=EcoRI Rsite2=XhoI Normal lung tissue from a 72
            year old male. Cloned unidirectionally. Primer: Oligo dT. Average
            insert size: 1.0 kb; Uni-ZAP XR Vector; 5' adaptor sequence:
            5'-GAATTCGGCACGAG-3; 3' adaptor sequence:
            5'-CTCGAGTTTTTTTTTTTTTTTTT-3'.
ORGANISM   Homo sapiens
            Eucaryotae; Metazoa; Chordata; Vertebrata; Gnathostomata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 264)
AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Tan,F., TrevaSkis,E.,
            Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.
TITLE   WashU-Merck EST Project
JOURNAL   Unpublished (1995)

---

22   62   3.0   R16276   ye48d02.r1 Homo sapie 1.18e-56
c   23   62   3.0   T92863   ye27f01.r1 Homo sapie 1.18e-56
c   24   53   2.6   T86143   yd84f01.s1 Homo sapie 1.41e-42
    25   38   1.8   R09732   yf27a12.s1 Homo sapie 9.86e-21
    26   37   1.8   R43932   yg22d07.s1 Homo sapie 2.23e-19
    27   37   1.8   R54590   yg81h09.s1 Homo sapie 2.23e-19
    28   37   1.8   T47099   yb52c08.s1 Homo sapie 2.23e-19
c   29   37   1.8   T07122   EST05011 Homo sapiens 2.23e-19
c   30   37   1.8   R06912   yf12g05.s1 Homo sapie 2.23e-19
c   31   36   1.7   T71627   yd36f08.s1 Homo sapie 4.87e-18
    32   36   1.7   T80892   yd26h11.r1 Homo sapie 4.87e-18
c   33   35   1.7   T58750   yB80e05.s1 Homo sapie 1.02e-16
    34   34   1.6   T96616   ye51e04.s1 Homo sapie 2.05e-15
    35   34   1.6   T89995   ye13b10.r1 Homo sapie 2.05e-15
c   36   33   1.6   HSC05A082 H. sapiens partial cD 3.94e-14
    37   33   1.6   HSC2TC062 H. sapiens partial cD 3.94e-14
c   38   33   1.6   T91101   yd51c05.s1 Homo sapie 3.94e-14
    39   32   1.5   T07131   EST05020 Homo sapiens 7.22e-13
c   40   32   1.5   R63483   yi08f10.r1 Homo sapie 7.22e-13
    41   32   1.5   T97712   ye54e02.s1 Homo sapie 7.22e-13
    42   32   1.5   T57458   yb51h09.r1 Homo sapie 7.22e-13
    43   32   1.5   R50724   yj58d01.s1 Homo sapie 7.22e-13
    44   32   1.5   HSAAAAVGP H. sapiens putatively 7.22e-13
c   45   32   1.5   R51961   yj71h01.r1 Homo sapie 7.22e-13

                    ALIGNMENTS

RESULT   1
LOCUS   T98687   506 bp   mRNA   31-MAR-1995
DEFINITION ye61b01.r1 Homo sapiens cDNA clone 122185 5' similar to gb:M92934
            CONNECTIVE TISSUE GROWTH FACTOR PRECURSOR (HUMAN);.
ACCESSION   T98687
KEYWORDS   EST.
SOURCE   human clone=122185 library=Soares fetal liver spleen 1NFLS
            vector=pT7T3D (Pharmacia) with a modified polylinker host=DH10B
            (ampicillin resistant) primer=M13RP1 Rsite1=Pac I Rsite2=Eco RI
            Liver and spleen from a 20 week-post conception male fetus. 1st
            strand cDNA was primed with a Pac I - oligo(dT) primer [5'
            AACTGGAAGAATTAATTAAGATCTTTTTTTTTTTTTTTTT 3'], double-stranded
            cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac
            I and cloned into the Pac I and Eco RI sites of the modified pT7T3
            vector. Library went through one round of normalization. Library
            constructed by Bento Soares and M.Fatima Bonaldo.
ORGANISM   Homo sapiens
            Eucaryotae; Metazoa; Chordata; Vertebrata; Gnathostomata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 506)
AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
TITLE   The WashU-Merck EST Project
JOURNAL   Unpublished (1995)
COMMENT   Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            High qality sequence stops: 303

COMMENT
```
                 Contact: Wilson RK
                 WashU-Merck EST Project
                 Washington University School of Medicine
                 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
                 Tel: 314 286 1800
                 Fax: 314 286 1810
                 Email: est@watson.wustl.edu
                 High quality sequence stops: 249
                 Source: IMAGE Consortium, LLNL
                 This clone is available royalty-free through LLNL ; contact the
                 IMAGE Consortium (info@image.llnl.gov) for further information.

FEATURES         Location/Qualifiers
     source      1..264
                 /organism="Homo sapiens"
                 /clone="118142"
                 /note="human"

NCBI gi: 724226

BASE COUNT       73 a     71 c     63 g     56 t     1 others
ORIGIN

DB 56;   Score  249; Match 98.9%; QryMatch 12.0%; Pred. No. 0.00e+00;
Matches  261; Conservative  0; Mismatches  1; Indels  2; Gaps  2;

Db     1 aaatcccagcctatcaagtttgagctttctggctgccaccagcatgaagacataccgag  60
         ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   911 AAATCTCCAAGCCTATCAAGTTTGAGCTTTCTGGCTGCCACCAGCATGAAGACATACCGAG  970

Db    61 ctaaattctgtggagtatgtaccgacgcgatgctgcacccccacaagaaccaccaccc  120
         |||||||||||||||||||||||||||||| |||||||||||||| ||||||||||||||||
Qy   971 CTAAATTCTGTGGAGTATGTACCGACGCGCCGATGCTGCACCCCCCAGAACCACCACCC  1030

Db   121 tgccggtngagttcaagtgccctgacggcgaggcgaggcatgaggaagaacatgatgttcatca  180
         |||||| |||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Qy  1031 TGCCGGTGGAGTTCAAGTGCCCTGACGGCGAGGTCAGGCATGAGGAAGAACATGATGTTCATCA  1090

Db   181 agacctgtgctgccattacaactgtcccggagacaatgacatctttggaatgctgtac  240
         |||||||||| ||||||||||||||||||||||||||||||||||||| ||| |||||||
Qy  1091 AGACCTGTGCCTCCATTACAACTGTCCCGGAGACAATGACATCTTTG-AATCGCTGTAC  1149

Db   241 tacaggaagatgtacggagacat  264
         |||||| |||||||||||||||||
Qy  1150 TACAGG-AAGATGTACGGAGACAT  1172
```

RESULT   3
LOCUS   T86251   434 bp   mRNA   EST   17-MAR-1995
DEFINITION yd84f01.r1 Homo sapiens cDNA clone 114937 5' similar to gb:M92934
            CONNECTIVE TISSUE GROWTH FACTOR PRECURSOR (HUMAN);.
ACCESSION   T86251
KEYWORDS   EST.
SOURCE   human clone=114937 library=Soares fetal liver spleen 1NFLS
            vector=pT7T3D (Pharmacia) with a modified polylinker host=DH10B
            (ampicillin resistant) primer=M13RP1 Rsite1=Pac I Rsite2=Eco RI
            Liver and spleen from a 20 week-post conception male fetus. 1st
            strand cDNA was primed with a Pac I - oligo(dT) primer [5'
            AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTT 3'], double-stranded
            cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac
            I and cloned into the Pac I and Eco RI sites of the modified pT7T3
            vector. Library went through one round of normalization. Library
            constructed by Bento Soares and M.Fatima Bonaldo.
  ORGANISM Homo sapiens
            Eukaryotae; Metazoa; Chordata; Vertebrata; Gnathostomata; Mammalia;

COMMENT
```
REFERENCE        Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS          1 (bases 1 to 434)
                 Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
                 Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
                 Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
                 Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
                 Wilson,R.
TITLE            The WashU-Merck EST Project
JOURNAL          Unpublished (1995)
COMMENT
                 Contact: Wilson RK
                 WashU-Merck EST Project
                 Washington University School of Medicine
                 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
                 Tel: 314 286 1800
                 Fax: 314 286 1810
                 Email: est@watson.wustl.edu
                 High quality sequence stops: 301
                 Source: IMAGE Consortium, LLNL
                 This clone is available royalty-free through LLNL ; contact the
                 IMAGE Consortium (info@image.llnl.gov) for further information.

FEATURES         Location/Qualifiers
     source      1..434
                 /organism="Homo sapiens"
                 /clone="114937"
                 /note="human"

NCBI gi: 714603

BASE COUNT       117 a     81 c    111 g    119 t     6 others
ORIGIN

DB 55;   Score  245; Match 94.7%; QryMatch 11.8%; Pred. No. 0.00e+00;
Matches  302; Conservative  0; Mismatches  9; Indels  8; Gaps  6;

Db     2 aatagtctatcaacccagacacctgtttgaagaatgttaagacttgacagtggaactac  61
         |||||||||| |||  |||||||| ||| ||||||||||||||||||||||||||||||||
Qy  1343 AATAGTCTATCTTCCCCAGACACTGGTTGAAGAATGTTAAGACTTGACAGTGGAACTAC  1402

Db    62 attagtacacagcaccagaatgtatattaaggtgtggctttaggagcagtgggagggtac  121
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1403 ATTAGTACACAGCACCAGAATGTATATTAAGGTGTGGCTTTAGGAGCAGTGGGAGGGTAC  1462

Db   122 cagcagaaaggttagtatcatcagatagcatcttatacgagtaatatgctgctatttga  181
         | ||   ||||||||||||||||||| ||  ||||||||||||||||| |||||||||||||
Qy  1463 CGGCCC---GGTTAGTATCATCAGATCGACTCTTATACGAGTAATAATGCCTGCTATTTGA  1519

Db   182 agtgtaattgagaagagaaaatttttagcgtgctcactgacctgctctagcctcagtgaca  241
         |||||||||||||| ||||||||||||||||||||||||||||| || |||| ||||||||||
Qy  1520 AGTGTAATTGAGAAGGAGAAAATTTTAGCGTGCTCACTGACCTGCCTGTAGCCCCAGTGACA  1579

Db   242 gctaggatggtggcattctccagccatccaaggagactgagtcaagttgttccnttaagtc  301
         |||||||||   ||||||||||||||| ||| ||||||||||||||||||||||| |||||||
Qy  1580 GCTAGGATGTG-CATTCTCCAGCCATC-AAG-AGACTGAGTCAAGTTGTTCC-TTAAGTC  1635

Db   302 aggaacagcagactcagct  320
         || ||||||||||||| |||
Qy  1636 AG-AACAGCAGACTCACCT  1653
```

RESULT   4
LOCUS   R37906   345 bp   mRNA   EST   04-MAY-1995
DEFINITION yh97e04.r1 Homo sapiens cDNA clone 137694 5' similar to gb:M92934
            CONNECTIVE TISSUE GROWTH FACTOR PRECURSOR (HUMAN);.
ACCESSION   R37906

Qy 1390 ACAGTGG-AACTACATTAGT-ACACAG-CACCAG-AATGTATATTAAGGTGTG 1438
        ||||||| |||||||||||| |||||| ||||| |||||| ||||| |||||| ||

RESULT 5
LOCUS   T98574   420 bp   mRNA   EST   31-MAR-1995
DEFINITION ye60a12.r1 Homo sapiens cDNA clone 122110 5' similar to gb:M92934
            CONNECTIVE TISSUE GROWTH FACTOR PRECURSOR (HUMAN);.
ACCESSION   T98574
KEYWORDS   EST.
SOURCE   human clone=122110 library=Soares fetal liver spleen 1NFLS
            vector=pT7T3D (Pharmacia) with a modified polylinker host=DH10B
            (ampicillin resistant) primer=M13RP1 Rsite1=Pac I Rsite2=Eco RI
            Liver and spleen from a 20 week-post conception male fetus. 1st
            strand cDNA was primed with a Pac I - oligo(dT) primer [5'
            AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTTTT 3'], double-stranded
            cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac
            I and cloned into the Pac I and Eco RI sites of the modified pT7T3
            vector. Library went through one round of normalization. Library
            constructed by Bento Soares and M.Fatima Bonaldo.
ORGANISM   Homo sapiens
            Eucaryotae; Metazoa; Chordata; Vertebrata; Gnathostomata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 420)
AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
TITLE   The WashU-Merck EST Project
JOURNAL   Unpublished (1995)
COMMENT   Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            High quality sequence stops: 232
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.

FEATURES   Location/Qualifiers
source   1..420
            /organism="Homo sapiens"
            /clone="122110"
            /note="human"
BASE COUNT   106 a   99 c   111 g   96 t   8 others
ORIGIN

NCBI gi: 748311

DB 58; Score   209; Match 95.6%; QryMatch 10.1%; Pred. No. 2.67e-317;
Matches 259; Conservative   0; Mismatches   4; Indels   8; Gaps   7;

Db   12 gacaacggcctcctgccaggctagagagaagcagaac-gcctgtgcatggtcaggccttgcga- 69
         |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Qy   799 GACAACGGCCTCCTGCCAGGCTAGAGAAGCAGCAGCCGCCGCCCTGTGCATGGTCAGGCCCTTGCGAA 858

Db   70 -cngacctggaagagagaagaacattaagaaggcaaaaaagtgcaaaaagtcgatccgtactcccaaaatttcc 128
         | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Qy   859 GCTGACCTGGAAGAGAGAACATTAAGAAGGGCAAAAAGTGCATCCGACTCCCAAAATCTCC 918

KEYWORDS   EST.
SOURCE   human clone=137694 library=Soares placenta Nb2HP vector=pT7T3D
            (Pharmacia) with a modified polylinker host=DH10B (ampicillin
            resistant) primer=M13RP1 Rsite1=Not I Rsite2=Eco RI Female placenta
            obtained at birth (full term). 1st strand cDNA was primed with a
            Not I - oligo(dT) primer [5'
            AACTGGAAGAATTCCGCGCCGCCAGGAATTTTTTTTTTTTTTT 3'], double-stranded
            cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not
            I and cloned into the Not I and Eco RI sites of the modified pT7T3
            vector. Library went through one round of normalization. Library
            constructed by Bento Soares and M.Fatima Bonaldo.
ORGANISM   Homo sapiens
            Eucaryotae; Metazoa; Chordata; Vertebrata; Gnathostomata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 345)
AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
TITLE   The WashU-Merck EST Project
JOURNAL   Unpublished (1995)
COMMENT   Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            High quality sequence stops: 262
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.

FEATURES   Location/Qualifiers
source   1..345
            /organism="Homo sapiens"
            /clone="137694"
            /note="human"
BASE COUNT   108 a   62 c   83 g   90 t   2 others
ORIGIN

NCBI gi: 795362

DB 25; Score   235; Match 95.2%; QryMatch 11.4%; Pred. No. 0.00e+00;
Matches 279; Conservative   0; Mismatches   0; Indels   6; Gaps   6;

Db   1 caggaagatgtacgcagacatggcatgaagccagagagtgagagacattaactcattaga 60
         ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Qy 1152 CAGGAAGATGTACGGAGACATGGCATGAAGCCAGAGAGTGAGAGACATTAACTCATTAGA 1211

Db   61 ctggaacttgaactgattcacatctcattttccgtaaaatgatttcagtagcacaagt 120
         |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Qy 1212 CTGGAACTTGAACTGATTCACATCTCATTTTTCCGTAAAATGATTTCAGTAGCACAAGT 1271

Db   121 tatttaaatctgttttctaactggggaaaagattcccacccaattcaaaacattgtgc 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1272 TATTTAAATCTGTTTTCTAACTGGGGAAAAGATTCCCACCCAATTCAAAACATTGTGC 1331

Db   181 catgtcaaacaaatagtctatncaacccagacactggttgaaggaatgtaagacttg 240
         |||||||||||||||||||||| |||||||||||||||||||||| |||||||||||
Qy 1332 CATGTCAAACAAATAGTCTAT-CTTCCCCAGACACTGGTTGAAG-AATGTTAAGACTTG 1389

Db   241 acagtgggaactacattagtcacaggcaccaggaatgtnttttaaggggtg 293

```
Db     1 acaacggctctgcaggctagagaagagagagccgcctgtgcatggtcaggccttgcgaag 60
         ||||||||| ||| ||||||||||| || ||||| || ||||| ||| || |||||||||
Qy   800 ACAACGGCTCCTCCCAGGCTAGAGAACGCAGAAGCCAGACCCGCCTGTTGCAGTGTCAGGCCTTGCGAAG 859

Db    61 tgaacctggagagaacattaagaaggccaaaagtgcatccgtactcccaaatttcca 120
         |||||||||||||||||||||||| ||| |||||||||||||||||||||||| |||
Qy   860 CTGACCTGGAGAGAACATTAAGAGGGCAAAAAGTGCATCCGTACTCCCAAAATCTCCA 919

Db   121 agcctatcaagtttgagctttctggctgcaccagcagcatgaaggacataccgagctaaattc 180
         ||||||||||||||||||||||||||||||||||||  |||||| |||||||||||||||||
Qy   920 AGCCTATCAAGTTTGAGCTTTCTGGCTGCACCAGCATGAAG-ACATACCGAGCTAAATTC 978

Db   181 tgtgggagtgatgtaccgacggccgatgctgtcacccctttacaggaaccacccctgccg 240
         |||||  ||| |||||||||||||||||||||    ||||  |||| |||||| ||||||
Qy   979 TGTGG-AGTATGTACCGACGGCCGATGCTGCTGCACCCCCC-ACAG-AACCACCACCCTGCCG 1035

Db   241 gtgg 244
         ||||
Qy  1036 GTGG 1039

RESULT      7
LOCUS       T49714       357 bp     mRNA     EST        08-FEB-1995
DEFINITION  ya80a10.r1 Homo sapiens cDNA clone 67962 5' similar to gb:M92934
            CONNECTIVE TISSUE GROWTH FACTOR PRECURSOR (HUMAN).
ACCESSION   T49714
KEYWORDS    EST.
SOURCE      human clone=67962 library=Stratagene ovary (#937217)
            vector=Bluescript SK host=SOLR cells (kanamycin resistant)
            primer=M13RP1 Rsite1=EcoRI Rsite2=XhoI Cloned unidirectionally.
            Primer: Oligo dT. Total ovary tissue, normal 49 year old caucasian
            female. Average insert size: 0.8 kb; Uni-ZAP XR Vector; 5' adaptor
            sequence: 5'-GAATTCGGCACGAG-3'; 3' adaptor sequence:
            5'-CTCGAGTTTTTTTTTTTTTTTTT-3'.
  ORGANISM  Homo sapiens
            Eucaryotae; Metazoa; Chordata; Vertebrata; Gnathostomata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 357)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Tan,F., Trevaskis,E.,
            Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Other ESTs: ya80a10.s1
            Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            High qality sequence stops: 254
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.

NCBI gi: 651574
FEATURES            Location/Qualifiers
     source         1..357
                    /organism="Homo sapiens"
                    /clone="67962"
                    /note="human"
```

```
Db   129 aagcctatcaagtttgagctttctggctgcaccagcagcatgaaggacataccgagctaaattc 188
         ||||||||||||||||||||||||||||||||||||  |||||| |||||||||||||||||
Qy   919 AGCCTATCAAGTTTGAGCTTTCTGGCTGCACCAGCATGAAG-ACATACCGAGCTAAATTC 978

Db   189 tgtgggagtgatgtaccgacggccgatgctgtcacccctttacaggaaccacccctgccgg 248
         |||||  ||| |||||||||||||||||||||    ||||  |||| |||||| |||||||
Qy   979 TGTGG-AGTATGTACCGACGGCCGATGCTGCTGCACCCCCC-ACAG-AACCACCACCCTGCCGG 1036

Db   249 tggagtttcaagttgcccttgacggcgaggt 279
         |||||||  |||| ||||| ||||||  |||
Qy  1037 TGGAGTT-CAAGT-GCCCT-GACGGCCAGGT 1064

RESULT      6
LOCUS       R06961       376 bp     mRNA     EST        05-APR-1995
DEFINITION  yf12g05.r1 Homo sapiens cDNA clone 126680 5' similar to gb:M92934
            CONNECTIVE TISSUE GROWTH FACTOR PRECURSOR (HUMAN);.
ACCESSION   R06961
KEYWORDS    EST.
SOURCE      human clone=126680 library=Soares fetal liver spleen 1NFLS
            vector=pT7T3D (Pharmacia) with a modified polylinker host=DH10B
            (ampicillin resistant) primer=M13RP1 Rsite1=Pac I Rsite2=Eco RI
            Liver and spleen from a 20 week-post conception male fetus. 1st
            strand cDNA was primed with a Pac I - oligo(dT) primer [5'
            AACTCGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTT 3'], double-stranded
            cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac
            I and cloned into the Pac I and Eco RI sites of the modified pT7T3
            vector. Library went through one round of normalization. Library
            constructed by Bento Soares and M.Fatima Bonaldo.
  ORGANISM  Homo sapiens
            Eucaryotae; Metazoa; Chordata; Vertebrata; Gnathostomata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 376)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            High qality sequence stops: 235
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.

NCBI gi: 758884
FEATURES            Location/Qualifiers
     source         1..376
                    /organism="Homo sapiens"
                    /clone="126680"
                    /note="human"

BASE COUNT     95 a     93 c     98 g     87 t        3 others
ORIGIN

DB 16;  Score   206;  Match 96.3%;  QryMatch  9.9%;  Pred. No. 9.94e-312;
Matches 235;  Conservative    0;  Mismatches    5;  Indels    4;  Gaps    4;
```

```
BASE COUNT    89 a    53 c    73 g    113 t    3 others
ORIGIN                          /note="human"

DB 56;  Score    187;  Match 99.5%;  QryMatch 9.0%;  Pred. No. 1.11e-276;
Matches  193;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Db     2  tatttcaatgttagcctcaatttctgaaccaccatagtagaatgtaaagcttgtctgatc 61
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1878  TATTTCAATGTTAGCCTCAATTTCTGAACACCATAGTAGAATGTAAAGCTTGTCTGATC 1937

Db    62  gttcaaagcatgaaatggtactatatggaaattctgctcagatagaatgacagtccgt 121
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1938  GTTCAAAGCATGAAATGGTACTATATGGAAATTCTGCTCAGATAGAATGACAGTCCGT 1997

Db   122  caaaacagattgtttgcaaagggagggaggcatcagtgtccttggcaggctgatttctaggta 181
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1998  CAAAACAGATTGTTTGCAAAGGGAGGCAGGCATCAGTGTC-TTGGCAGGCTGATTTCTAGGTA 2056

Db   182  ggaaatgtggtagc 195
          ||||||||||||||
Qy  2057  GGAAATGTGGTAGC 2070

RESULT    9
LOCUS         T65929          410 bp    mRNA          EST          20-FEB-1995
DEFINITION    yc24e08.s1 Homo sapiens cDNA clone 81638 3' similar to gb:M92934
              CONNECTIVE TISSUE GROWTH FACTOR PRECURSOR (HUMAN);.
ACCESSION     T65929
KEYWORDS      EST.
SOURCE        human clone=81638 library=Stratagene lung (#937210)
              vector=pBluescript SK- host=SOLR cells (kanamycin resistant)
              primer=21m13 Rsite1=EcoRI Rsite2=XhoI Normal lung tissue from a 72
              year old male. Cloned unidirectionally. Primer: Oligo dT. Average
              insert size: 1.0 kb; Uni-ZAP XR Vector; 5' adaptor sequence:
              5'-GAATTCGGCACGAG-3'; 3' adaptor sequence:
              5'-CTCGAGTTTTTTTTTTTTTTTTT-3'.
ORGANISM      Homo sapiens
              Eucaryotae; Metazoa; Chordata; Vertebrata; Gnathostomata; Mammalia;
              Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE     1  (bases 1 to 410)
AUTHORS       Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
              Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
              Parsons,J., Rifkin,L., Rohlfing,T., Tan,F., Trevaskis,E.,
              Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.
TITLE         WashU-Merck EST Project
JOURNAL       Unpublished (1995)
COMMENT       Contact: Wilson RK
              WashU-Merck EST Project
              Washington University School of Medicine
              4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
              Tel: 314 286 1800
              Fax: 314 286 1810
              Email: est@watson.wustl.edu
              High qality sequence stops: 377
              Source: IMAGE Consortium, LLNL
              This clone is available royalty-free through LLNL ; contact the
              IMAGE Consortium (info@image.llnl.gov) for further information.

NCBI gi: 674974
FEATURES             Location/Qualifiers
source               1..410
                     /organism="Homo sapiens"
```

```
BASE COUNT    96 a    54 c    83 g    124 t
ORIGIN

DB 45;  Score    198;  Match 99.5%;  QryMatch 9.5%;  Pred. No. 6.14e-297;
Matches  204;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Db     1  taatgctttgatattcaatgttagcctcaatttctgaacaccatagtagaatgtaaag 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1867  TAATGCTTTGATATTTCAATGTTAGCCTCAATTTCTGAACACCATAGTAGAATGTAAAG 1926

Db    61  cttgtctgatcgttcaaagcatgaaatggatactatatggaaattctgctcagatagaa 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1927  CTTGTCTGATCGTTCAAAGCATGAAATGGATACTATATGGAAATTCTGCTCAGATAGAA 1986

Db   121  tgacagtccgtcaaaacagattgtttgcaaagggaggcatcagtgtccttggcaggctg 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1987  TGACAGTCCGTCAAAACAGATTGTTTGCAAAGGGAGGCAGGCATCAGTGTC-TTGGCAGGCTG 2045

Db   181  attttcaggtaggaaatgtggtagc 205
          |||||||||||||||||||
Qy  2046  ATTTCTAGGTAGGAAATGTGGTAGC 2070

RESULT    8
LOCUS         T92598          331 bp    mRNA          EST          22-MAR-1995
DEFINITION    ye26f01.s1 Homo sapiens cDNA clone 118873 3' similar to gb:M92934
              CONNECTIVE TISSUE GROWTH FACTOR PRECURSOR (HUMAN);.
ACCESSION     T92598
KEYWORDS      EST.
SOURCE        human clone=118873 library=Stratagene lung (#937210)
              vector=pBluescript SK- host=SOLR cells (kanamycin resistant)
              primer=21m13 Rsite1=EcoRI Rsite2=XhoI Normal lung tissue from a 72
              year old male. Cloned unidirectionally. Primer: Oligo dT. Average
              insert size: 1.0 kb; Uni-ZAP XR Vector; 5' adaptor sequence:
              5'-GAATTCGGCACGAG-3'; 3' adaptor sequence:
              5'-CTCGAGTTTTTTTTTTTTTTTTT-3'.
ORGANISM      Homo sapiens
              Eucaryotae; Metazoa; Chordata; Vertebrata; Gnathostomata; Mammalia;
              Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE     1  (bases 1 to 331)
AUTHORS       Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
              Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
              Parsons,J., Rifkin,L., Rohlfing,T., Tan,F., Trevaskis,E.,
              Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.
TITLE         WashU-Merck EST Project
JOURNAL       Unpublished (1995)
COMMENT       Contact: Wilson RK
              WashU-Merck EST Project
              Washington University School of Medicine
              4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
              Tel: 314 286 1800
              Fax: 314 286 1810
              Email: est@watson.wustl.edu
              High quality sequence stops: 172
              Source: IMAGE Consortium, LLNL
              This clone is available royalty-free through LLNL ; contact the
              IMAGE Consortium (info@image.llnl.gov) for further information.

NCBI gi: 724511
FEATURES             Location/Qualifiers
source               1..331
                     /organism="Homo sapiens"
                     /clone="118873"
```

/organism="Homo sapiens"
/clone="67962"
/note="human"

BASE COUNT   133 a   82 c   65 g   118 t   10 others
ORIGIN

DB 45; Score 169; Match 94.0%; QryMatch 8.1%; Pred. No. 1.16e-243;
Matches 189; Conservative 0; Mismatches 9; Indels 3; Gaps 3;

Db   204 gctaccacatttcctacctagnaatcagcctgccaaggacaactgatgcctccccttgca 263
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Cp 2070 GCTACCACATTTCCTACCTAGAAATCAGCCTGCCAAG-ACACTGATGCCTCCCCTTTGCA 2012

Db   264 aacaatctgtttgacggactgtcattctatctgngcagantttccatatangnatccat 323
         |||||||||||||||||||||||||||||||||||||||||| |||||||
Cp 2011 AACAATCTGTTTGACGGACTGTCATTCTATCTGAGCAGAATTTCCATATAAGTATCCAT 1952

Db   324 ttcatgctttgaacgntcagncaagcttacattctacctatgggtcagnaaattgag 383
         ||||||||||||||||| |||| |||||||||||||||||||||||| |||||||
Cp 1951 TTCATGCTTTGAACGATCAGCAAGCTTACATTCTACCTATGCTTCAG-AAATTGAG 1893

Db   384 gctaacattgaaatatccaaa 404
         ||| |||||||| ||||||| ||
Cp 1892 GCTAACATTG-AAATATCAAA 1873

RESULT   11
LOCUS   T53387   365 bp   mRNA   EST   08-FEB-1995
DEFINITION   ya88f03.r1 Homo sapiens cDNA clone 68765 5' similar to gb:M92934
             CONNECTIVE TISSUE GROWTH FACTOR PRECURSOR (HUMAN).
ACCESSION   T53387
KEYWORDS   EST.
SOURCE   human clone=68765 library=Stratagene placenta (#937225)
             vector=pBluescript SK- host=SOLR cells (kanamycin resistant)
             primer=M13RP1 Rsite1=EcoRI Rsite2=XhoI Placental tissue from a
             Caucasian male. Cloned unidirectionally. Primer: Oligo dT. Average
             insert size: 1.2 kb; Uni-ZAP XR Vector; 5' adaptor sequence:
             5'-GAATTCGGCACGAG-3'; 3' adaptor sequence:
             5'-CTCGAGTTTTTTTTTTTTTTTT-3'.
ORGANISM   Homo sapiens
             Eucaryotae; Metazoa; Chordata; Vertebrata; Gnathostomata; Mammalia;
             Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 365)
AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
             Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
             Parsons,J., Rifkin,L., Rohlfing,T., Tan,F., Trevaskis,E.,
             Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.
TITLE   WashU-Merck EST Project
JOURNAL   Unpublished (1995)
COMMENT   Other_ESTs: ya88f03.s1
             Contact: Wilson RK
             WashU-Merck EST Project
             Washington University School of Medicine
             4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
             Tel: 314 286 1800
             Fax: 314 286 1810
             Email: est@watson.wustl.edu
             High qality sequence stops: 311
             Source: IMAGE Consortium, LLNL
             This clone is available royalty-free through LLNL ; contact the
             IMAGE Consortium (info@image.llnl.gov) for further information.

             NCBI gi: 655247

FEATURES   Location/Qualifiers

/clone="81638"
/note="human"

BASE COUNT   143 a   82 c   68 g   114 t   3 others
ORIGIN

DB 49; Score 172; Match 96.7%; QryMatch 8.3%; Pred. No. 3.81e-249;
Matches 208; Conservative 0; Mismatches 1; Indels 6; Gaps 6;

Db   196 gctaccacattcctacctagaaatcagcctgccaaggacaactgatgcctccccctttgca 255
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Cp 2070 GCTACCACATTCCTACCTAGAAATCAGCCTGCCAAG-ACACTGATGCCTCCCCTTTGCA 2012

Db   256 aacaatctgtttgacggactgtcattctatctgagcagaattccatataagtatccat 315
         ||||||||||||||||||||||||||||||||||||||||||||||||
Cp 2011 AACAATCTGTTTGACGGACTGTCATTCTATCTGAGCAGAATTTCCATATAAGTATCCAT 1952

Db   316 ttcatgctttgaacgatcagacaagcttacattctacctatgnggttcaggaaattga 375
         |||||||||||||||||||||||||||||||||| |||||||||| ||||||||
Cp 1951 TTCATGCTTTGAACGATCAGCAAGCTTACATTCTACCTATGCTG-TTCAG-AAATTGA 1894

Db   376 gggctaacattgaaatatcaaaggcattaaaaac 410
         || |||||||| |||||| |||||||
Cp 1893 GG-CTAACATTG-AAATATCAAAG-CATTAAAAAC 1862

RESULT   10
LOCUS   T49715   408 bp   mRNA   EST   08-FEB-1995
DEFINITION   ya80a10.s1 Homo sapiens cDNA clone 67962 3' similar to gb:M92934
             CONNECTIVE TISSUE GROWTH FACTOR PRECURSOR (HUMAN).
ACCESSION   T49715
KEYWORDS   EST.
SOURCE   human clone=67962 library=Stratagene ovary (#937217)
             vector=Bluescript SK host=SOLR cells (kanamycin resistant)
             primer=21ml3 Rsite1=EcoRI Rsite2=XhoI Cloned unidirectionally.
             Primer: Oligo dT. Total ovary tissue, normal 49 year old caucasian
             female. Average insert size: 0.8 kb; Uni-ZAP XR Vector; 5' adaptor
             sequence: 5'-GAATTCGGCACGAG-3'; 3' adaptor sequence:
             5'-CTCGAGTTTTTTTTTTTTTTTT-3'.
ORGANISM   Homo sapiens
             Eucaryotae; Metazoa; Chordata; Vertebrata; Gnathostomata; Mammalia;
             Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 408)
AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
             Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
             Parsons,J., Rifkin,L., Rohlfing,T., Tan,F., Trevaskis,E.,
             Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.
TITLE   WashU-Merck EST Project
JOURNAL   Unpublished (1995)
COMMENT   Other_ESTs: ya80a10.r1
             Contact: Wilson RK
             WashU-Merck EST Project
             Washington University School of Medicine
             4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
             Tel: 314 286 1800
             Fax: 314 286 1810
             Email: est@watson.wustl.edu
             High qality sequence stops: 364
             Source: IMAGE Consortium, LLNL
             This clone is available royalty-free through LLNL ; contact the
             IMAGE Consortium (info@image.llnl.gov) for further information.

             NCBI gi: 651575

FEATURES   Location/Qualifiers
source   1..408

```
                /organism="Homo sapiens"
                /clone="69481"
                /note="human"
BASE COUNT    29 a    43 c    42 g    51 t    40 others
ORIGIN

DB 46;  Score   157;  Match 78.8%;  QryMatch 7.6%;  Pred. No. 8.41e-222;
Matches 160;  Conservative  0;  Mismatches 43;  Indels  0;  Gaps  0;

Db       3  tatgntcnctntttgnngtgtnattaaganggannatttnngcntnctcactgacctgcc   62
            |||| || ||||| || |||| ||| || ||||  || | ||||||||||||||
Qy    1504  TATGCCTGCTATTTGAGTGTAATTGAGAAGGAAAATTTTAGCCTGCTCACTGACCTGCC  1563

Db      63  tntagcccntgacngctnggannntgcattctncnncctcnnggactggtgcnagtt  122
            | |||||| |||||| ||| ||| || |  || ||||||||| |||
Qy    1564  TGTAGCCCCAGTGACAGCTAGGATGTGCATTCTCCAGCCATCAAGAGACTGAGTCAAGTT  1623

Db     123  gttccttnagncngaacngnagactcngcnctgacnttcgnntgacacngttca  182
            ||||||| ||| || ||  |||| || |||||| |||  ||||| |||||
Qy    1624  GTTCCTTAAGTCAGAACAGCAGACTCAGCCTGACATTCTGATTCGAATGACACTGTTCA  1683

Db     183  gggnncggatcctgtcgattag  205
            || |||||||||||||||||
Qy    1684  GGAATCGGAATCCTGTCGATTAG  1706
```

```
RESULT  13
LOCUS       T27642    404 bp   mRNA    EST    04-JAN-1995
DEFINITION  EST10485 Homo sapiens cDNA 3' end similar to connective tissue
            growth factor (HT:1399).
ACCESSION   T27642
KEYWORDS    EST.
SOURCE      human library=Human Adipose tissue primer=M13-21.
ORGANISM    Homo sapiens
            Eucaryotae; Metazoa; Chordata; Vertebrata; Gnathostomata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 404)
AUTHORS     Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
            Bult,C.J., Lee,N., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
            White,O., Sutton,G., Blake,J.A., Brandon,R.C., Chiu,M.-W.,
            Clayton,R.A., Cline,R.T., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
            FitzGerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S.M.,
            Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle Jr,P.S.,
            Kelley,J.M., Klimek,K.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M.,
            Merrick,J.M., Moreno-Palanques,R.F., McDonald,L.A., Nguyen,D.T.,
            Pellegrino,S.M., Phillips,C.A., Ryder,S.E., Scott,J.L.,
            Saudek,D.M., Shirley,R., Small,K.V., Spriggs,T.A., Utterback,T.R.,
            Weidman,J.F., Li,Y., Bednarik,D.P., Cao,L., Cepeda,M.A.,
            Coleman,T.A., Collins,E.-J., Dimke,D., Feng,P., Ferrie,A.,
            Fischer,C., Hastings,G.A., He,W.-W., Hu,J.-S., Greene,J.M.,
            Gruber,J., Hudson,P., Kim,A., Kozak,D.L., Kunsch,C., Ji,H., Li,H.,
            Meissner,P.S., Olsen,H., Raymond,L., Wei,Y.-F., Wing,J., Xu,C.,
            Yu,G.-L., Ruben,S.M., Dillon,P.J., Fannon,M.R., Rosen,C.A.,
            Haseltine,W.A., Fields,C., Fraser,C.M. and Venter,J.C.
TITLE       Initial Assessment of Human Gene Diversity and Expression Patterns
            Based Upon 52 Million Basepairs of cDNA Sequence
JOURNAL     Unpublished (1995)
COMMENT     Other ESTs: EST10484 THC23588
            Contact: Venter, JC
            The Institute for Genomic Research
            932 Clopper Rd, Gaithersburg, MD 20878
            Tel: 3018699056
            Fax: 3018699423
            Email: tdbinfo@tdb.tigr.org
```

```
     source     1..365
                /organism="Homo sapiens"
                /clone="6876S"
                /note="human"
BASE COUNT    108 a    54 c    78 g    125 t
ORIGIN

DB 46;  Score   167;  Match 99.4%;  QryMatch 8.0%;  Pred. No. 5.21e-240;
Matches 173;  Conservative  0;  Mismatches  0;  Indels  1;  Gaps  1;

Db       4  tttctgaacaccataggtagaatgtaaagcttgtctgatcgttcaaagcatgaaatggat   63
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    1898  TTTCTGAACACCATAGGTAGAATGTAAAGCTTGTCTGATCGTTCAAAGCATGAAATGGAT  1957

Db      64  acttatatggaaattctgtcagatagaatgacagtccgtcaaaacagattgttgcaaa  123
            |||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Qy    1958  ACTTATATGGAAATTCTGTCAGATAGAATAGACAGTCCGTCAAAACAGATTGTTGCAAA  2017

Db     124  gggaggcatcagtgtccttggcaggctgattctaggtaggaaatgtggtagc  177
            |||||||||||||||| |||||||||||||||             |||||||||||||
Qy    2018  GGGAGGCATCAGTGTC-TTGGCAGGCTGATTCTGATTCTGATTTCTAGGTAGGAAATGTGGTAGC  2070
```

```
RESULT  12
LOCUS       T53524    205 bp   mRNA    EST    08-FEB-1995
DEFINITION  ya96b01.r1 Homo sapiens cDNA clone 69481 5' similar to gb:M92934
            CONNECTIVE TISSUE GROWTH FACTOR PRECURSOR (HUMAN).
ACCESSION   T53524
KEYWORDS    EST.
SOURCE      human clone=69481 library=Stratagene placenta (#937225)
            vector=pBluescript SK- host=SOLR cells (kanamycin resistant)
            primer=M3RP1 Rsite1=EcoRI Rsite2=XhoI Placental tissue from a
            Caucasian male. Cloned unidirectionally. Primer: Oligo dT. Average
            insert size: 1.2 kb; Uni-ZAP XR Vector; 5' adaptor sequence:
            5'-GAATTCGGCACGAG-3'; 3' adaptor sequence:
            5'-CTCGAGTTTTTTTTTTTTTTTTTT-3'.
ORGANISM    Homo sapiens
            Eucaryotae; Metazoa; Chordata; Vertebrata; Gnathostomata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 205)
AUTHORS     Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Tan,F., Trevaskis,E.,
            Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.
TITLE       WashU-Merck EST Project
JOURNAL     Unpublished (1995)
COMMENT     Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            High qality sequence starts: 1
            High qality sequence stops: 1
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Trace considered overall poor quality.
FEATURES             Location/Qualifiers
     source          NCBI gi: 655384
                     1..205
```

For clone availability, additional sequence and expression information related to this EST, please contact the TIGR Database (tdbinfo@tdb.tigr.org).

NCBI gi: 609740

```
FEATURES        Location/Qualifiers
     source     1..404
                /organism="Homo sapiens"
                /note="human"

BASE COUNT      140 a     82 c     67 g     113 t
ORIGIN
```

DB 39; Score 155; Match 95.3%; QryMatch 7.5%; Pred. No. 3.59e-218;
Matches 204; Conservative 0; Mismatches 2; Indels 8; Gaps 8;

```
Db   191 gctaccacatttcctacctagaaatcagcctgccaagcacactgatgcctcccctttgca 250
         ||||||||||||||||||||||||||||||||||||  |||||||||||||||||||||||
Cp  2070 GCTACCACATTTCCTACCTAGAAATCAGCCCTGCCAAG-ACACTGATGCCTCCCCTTTGCA 2012

Db   251 aacaatcgtttttgacggactgtcattctatctgagcaggaatttccatataagtatcc 310
         |||||||||||| ||||| ||||||||||||||||||| ||||||||||||||||||||
Cp  2011 AACAATCGTTTTGACGG-ACTGTCATTCTATCTGAGCAG-AATTCCATATAAGTATCC 1954

Db   311 atttcatgctttgaacggtcagacaagctttacattctacctatgnggttcagaaattg 370
         |||||||||||||||||||||||||| |||||||||| |||||| ||||||||||||
Cp  1953 ATTTCATGCTTTGAACGGTCAGACAGACAAGCTTTACATTCCACCTATGGTG-TTCAGAAATTG 1895

Db   371 agggctaaccattggaaatatcaaaggcattaa 404
         ||| |||| |||| |||||||||||| ||||||
Cp  1894 AGG-CTAAC-ATTG-AAATATCAAAG-CATTAAA 1865
```

RESULT 14
```
LOCUS       T64503      140 bp    mRNA            EST         20-FEB-1995
DEFINITION  yc24e08.r1 Homo sapiens cDNA clone 81638 5' similar to gb:M92934
            CONNECTIVE TISSUE GROWTH FACTOR PRECURSOR (HUMAN);.
ACCESSION   T64503
KEYWORDS    EST.
SOURCE      human clone=81638 library=Stratagene lung (#937210)
            vector=pBluescript SK- host=SOLR cells (kanamycin resistant)
            primer=M13RP1 Rsite1=EcoRI Rsite2=XhoI Normal lung tissue from a 72
            year old male. Cloned unidirectionally. Primer: Oligo dT. Average
            insert size: 1.0 kb; Uni-ZAP XR Vector; 5' adaptor sequence:
            5'-GAATTCGGCACGAG-3'; 3' adaptor sequence:
            5'-CTCGAGTTTTTTTTTTTTTTTT-3'.
ORGANISM    Homo sapiens
            Eucaryotae; Metazoa; Chordata; Vertebrata; Gnathostomata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 140)
AUTHORS     Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Tan,F., Trevaskis,E.,
            Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.
TITLE       WashU-Merck EST Project
JOURNAL     Unpublished (1995)
COMMENT     Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Source: IMAGE Consortium, LLNL
```

This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.

NCBI gi: 673548

```
FEATURES        Location/Qualifiers
     source     1..140
                /organism="Homo sapiens"
                /clone="81638"
                /note="human"

BASE COUNT      39 a     39 c     36 g     25 t     1 others
ORIGIN
```

DB 49; Score 139; Match 99.3%; QryMatch 6.7%; Pred. No. 2.86e-189;
Matches 139; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Db     1 tgaagacataccgagctaaatctgtggatgtgtaccgacggncgatgtgcaccccc 60
         |||||||||||||||||||| ||||||||| ||||||||||| ||||| |||||||||
Qy   956 TGAAGACATACCGAGCTAAATTCTGTGGAGTATCTACCGACGGCCGATGCTGCACCCCCC 1015

Db    61 acagaaccaccaccctgccgtggagttcaagtgccctgacggcgagtcatgaagaaga 120
         ||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Qy  1016 ACAGAACCACCACCCTGCCGTGGAGTTCAAGTCCCTGACGGCGAGTCATGAAGAAGA 1075

Db   121 acatgatgttcatcaagacc 140
         ||||||||||||||||||||
Qy  1076 ACATGATGTTCATCAAGACC 1095
```

RESULT 15
```
LOCUS       HHEA46W     326 bp    RNA             EST         18-AUG-1994
DEFINITION  H. sapiens partial cDNA sequence; transcribed sequence fragment.
ACCESSION   236725
KEYWORDS    partial cDNA sequence; transcribed sequence fragment.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Eumycota group;
            Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata;
            Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Choanata;
            Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates;
            Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 326)
AUTHORS     Genexpress.
TITLE       Direct Submission
JOURNAL     Submitted (12-AUG-1994) to the EMBL/GenBank/DDBJ databases.
            Genzentrum Muenchen, Laboratorium fuer molekulare Biologie, Am
            Klopferspitz 18a, 82152 Martinsried, Germany. E-mail:
            obermaier@vms.biochem.mpg.de
REFERENCE   2 (bases 1 to 326)
AUTHORS     Genexpress.
TITLE       The Genexpress cDNA program
JOURNAL     Unpublished
COMMENT     cloning vector is pBluescript SK(+);
            Genexpress library reference is A.
            automatic.
```

NCBI gi: 531575

```
FEATURES        Location/Qualifiers
     source     1..326
                /organism="Homo sapiens"
                /clone lib="Atrium cDNA library Human heart"

BASE COUNT      113 a     68 c     48 g     97 t
ORIGIN
```

DB 3; Score 124; Match 99.2%; QryMatch 6.0%; Pred. No. 1.88e-162;

```
Matches    130;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Db    196 gctaccacattcctacctagaaatcagcctgccaaggacactgatgctccccttgca 255
          ||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Cp   2070 GCTACCACATTTCCTACCTAGAAATCAGCCTGCCAAG-ACACTGATGCCTCCCTTTGCA 2012

Db    256 aacaatctgtttgacggactgtcattctatctgagcagaatttccatataagtatccat 315
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Cp   2011 AACAATCTGTTTGACGGACTGTCATTCTATCTGAGCAGAATTTCCATATAAGTATCCAT 1952

Db    316 ttcatgctttg 326
          |||||||||||
Cp   1951 TTCATGCTTTG 1941

Search completed: Fri Sep 22 21:57:17 1995
Job time : 425 secs.
```